LARGE DEVIATION PRINCIPLE FOR FUNCTIONAL LIMIT THEOREMS

by

ADINA OPRISAN

Presented to the Faculty of the Graduate School of The University of Texas at Arlington in Partial Fulfillment of the Requirements for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT ARLINGTON

August 2009
Copyright © by Adina Oprisan  2009
All Rights Reserved
This dissertation is dedicated to the memory of my father in law

Professor Gh. Oprisan
ACKNOWLEDGEMENTS

First and foremost, I would like to express my indebted gratitude to my adviser Prof. Andrzej Korzeniowski for his warm encouragement, thoughtful guidance and support and for sharing with me a lot of his expertise and research insight. Also, I would like to thank Prof. Han Chien-Pai, Prof. D.L. Hawkins, Prof. Jianzhong Su and Prof. Shan Sun-Mitchell for their kind agreement to serve on my committee.

I feel very privileged to have met the families of Professors Corduneanu and Dragan, who substituted my home family and I would like to thank them very much for all unforgettable moments spent together.

Among the friends I made as a graduate student, a special mention is due to Violeta Yurita and Silviu Minut. I shall cherish their friendship forever.

I owe so much thanks to my parents who were continuously supporting me throughout my life.

I am forever indebted to my husband Paul for his support and assistance in every step of my thesis. I would like to thank to my daughter Dana for being such an understanding child and for giving me unlimited happiness and strength to finish this thesis.

July 7, 2009
LARGE DEVIATION PRINCIPLE FOR FUNCTIONAL LIMIT THEOREMS

Adina Oprisan, Ph.D.
The University of Texas at Arlington, 2009

Supervising Professor: Andrzej Korzeniowski

We study a family of stochastic additive functionals of Markov processes with locally independent increments switched by jump Markov processes in an asymptotic split phase space. Based on an averaging limit theorem, we obtain a large deviation result for this stochastic evolutionary system using a weak convergence approach. Examples, including compound Poisson processes, illustrate cases in which the rate function is calculated in an explicit form. We prove also a large deviation principle for a class of empirical processes in \( C[0, \infty) \) associated with additive functionals of Markov processes that were shown to have a martingale decomposition. Functional almost everywhere central limit theorems are established and the large deviation results are derived.
# TABLE OF CONTENTS

ACKNOWLEDGEMENTS .................................................. iv

ABSTRACT ................................................................. v

Chapter Page

1. INTRODUCTION ....................................................... 1

2. GENERAL FRAMEWORK ............................................... 3
   2.1 Markov processes: notations and definitions ......................... 3
      2.1.1 Markov chains .................................................. 5
      2.1.2 Continuous-time Markov processes ............................ 8
      2.1.3 Diffusion processes ............................................ 13
      2.1.4 Processes with independent increments ........................ 14
      2.1.5 Processes with locally independent increments .................. 17
   2.2 Weak convergence of stochastic processes .......................... 18
      2.2.1 Martingale characterization of Markov processes ............. 18
      2.2.2 Reducible-invertible operators ................................ 21
      2.2.3 Perturbation of reducible-invertible operators ............... 23
      2.2.4 Weak convergence .............................................. 25
   2.3 Large deviation principle ........................................ 28

3. LARGE DEVIATION PRINCIPLE FOR AVERAGE LIMIT THEOREMS .... 38
   3.1 Average approximation ............................................ 39
      3.1.1 Stochastic additive functionals ............................... 39
      3.1.2 Phase merging principle ....................................... 41
   3.2 Large deviation principle for ergodic Markov processes ............ 54
CHAPTER 1
INTRODUCTION

The theory of large deviations concerns the asymptotic behavior of remote tails of sequences of probability distributions. It has applications in many areas including statistics, communication networks, information theory, statistical mechanics, risk-sensitive control and queuing systems.

The first rigorous large deviations results were obtained by Harald Cramér in the late 1930s, who applied them to model the insurance business. He gave a large deviation principle (LDP) result for a sum of i.i.d. random variables with the rate function expressed by a power series. A general abstract framework for the LDP was proposed by S.R.S. Varadhan in 1966. At that time, the only modern large deviation principles available were: Schilder (LDP for Brownian motion), Sanov (LDP for ergodic processes), and Freidlin and Wentzell (LDP for Itô diffusions) with some abstract foundation of LDP. A crucial step forward was achieved through a series of papers of Donsker and Varadhan, starting in the mid 1970s in which they developed a systematic large deviation theory for empirical measures in the i.i.d. and Markov cases. For his seminal contributions to the field, Varadhan has won the 2007 Abel Prize. A unified approach, including Non-Uniform and Hypoeliptic Diffusions is due to D.W. Stroock.

In this dissertation we establish several large deviations results for stochastic additive functionals that are strongly or weakly convergent under different scalings. In Chapter 2 we introduce the necessary notations, definitions and state basic facts and results related to our study. Furthermore, the fundamental objects such as Markov
chains, Markov processes and some of their sub-families, including their properties are discussed and the framework for our analysis is presented.

In Chapter 3 we present the notions of switching and switched processes via additive functionals of processes with locally independent increments subject to Markov switching in phase split and merging scheme. The method is based on a splitting of the phase space into disjoint classes $E_k, k \in V$ and further merging these classes into distinct states $k \in V$. Large deviation principle for phase merging principle and average approximation theorem are established. These results are published in [1]. For phase merging principle the large deviation result was obtained via the LDP considerations for ergodic processes. The large deviation result for average approximation theorem follows from a weak convergence approach. That is, we represent the normalized logarithms of the expectations as variational formulas that are interpreted as the minimal cost functions of associated stochastic optimal control problems.

In Chapter 4 we study additive functionals associated with functional central limit theorems using entirely different techniques than those used in Chapter 3. We establish functional almost everywhere central limit theorems and derive the large deviation principle through the rate function for Ornstein-Uhlenbeck processes and a martingale decomposition of the additive functionals. These results are published in [2].

Finally, in Chapter 5 we outline some ideas for future work, including the extension of the results obtained in Chapter 4 for additive functionals in continuous time and for stochastic additive functionals.
CHAPTER 2
GENERAL FRAMEWORK

2.1 Markov processes: notations and definitions

Let \((E, r)\) be a complete, separable metric space, and let \(\mathcal{E}\) be its Borel \(\sigma\)-algebra of all Borel subsets of \(E\). We will call the measurable space \((E, \mathcal{B}_E)\) a standard state space. The space \(\mathbb{R}^d\), with the Euclidian metric, is a complete, separable metric space, and define by \(\mathcal{B}_d\) its Borel \(\sigma\)-algebra.

The space where the trajectories of processes are considered is \(\mathcal{D}[0, \infty)\), the space of right continuous functions having left hand side limits. This embedded with Skorokhod metric becomes a complete, separable metric space. Let also \((C[0, \infty), \| \cdot \|)\) be the space of continuous functions with the sup-norm, \(\| x \| = \sup_{t \geq 0} |x(t)|, \ x \in C[0, \infty)\).

Let \(B(E)\) the Banach space of all bounded, Borel measurable functions. Endowed with the norm \(\| \varphi \| = \sup_{x \in E} |\varphi(x)|\), \((B(E), \| \cdot \|)\) is a Banach space.

Let \(\Omega\) the stochastic basis \((\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})\) where \((\mathcal{F}_t)_{t \geq 0}\) is a filtration of sub-\(\sigma\)-algebras of \(\mathcal{F}\), that is: \(\mathcal{F}_s \subseteq \mathcal{F}_t \subseteq \mathcal{F}\), for all \(s < t\) and \(t \geq 0\). The filtration is said to be complete if \(\mathcal{F}_0\) contains all the \(\mathbb{P}\)-null sets. Set \(\mathcal{F}_{t+} = \cap_{s > t} \mathcal{F}_s\), \(t \geq 0\). If for any \(t \geq 0\), \(\mathcal{F}_t = \mathcal{F}_{t+}\), then the filtration \(\mathcal{F}_t\), \(t \geq 0\) is said to be right-continuous. If a filtration is complete and right-continuous, we say that it satisfies the usual conditions.

A mapping \(T : \Omega \rightarrow [0, \infty]\), such that \(\{ T \leq t \} \in \mathcal{F}_t\) is called a stopping time. If \(T\) is a stopping time, we denote by \(\mathcal{F}_T\) the collection of all sets \(A \in \mathcal{F}\) such that \(A \cap \{ T \leq t \} \in \mathcal{F}_t\).
An \((E, \mathcal{B}_E)\)-valued stochastic process \(x(t), t \in I\), defined on \(\Omega\) is adapted to the filtration if for any \(t \in I\), \(x(t)\) is \(\mathcal{F}_t\)-measurable. The set of values \(E\) is said to be the state (phase) space of the process \(x(t), t \geq 0\).

Given a probability measure \(\mu\) on \((E, \mathcal{B}_E)\), we define the probability measures \(\mathbb{P}_\mu\), by
\[
\mathbb{P}_\mu(B) = \mu \mathbb{P}(B) = \int_E \mu(dx) \mathbb{P}_x(B), \quad x \in E, B \in \mathcal{F}
\]
where \(\mathbb{P}_x(B) := \mathbb{P}(B|x(0) = x)\). We denote by \(\mathbb{E}_\mu\) and \(\mathbb{E}_x\) the expectations corresponding respectively to \(\mathbb{P}_\mu\) and \(\mathbb{P}_x\).

A positive-valued function \(P(x, B), x \in E, B \in \mathcal{B}_E\) is called a Markov transition function or a Markov kernel if
1) for any fixed \(x \in E\), \(P(x, \cdot)\) is a probability measure on \((E, \mathcal{B}_E)\)
2) for any fixed \(B \in \mathcal{B}_E\), \(P(\cdot, B)\) is a Borel measurable function.

If \(P(x, E) \leq 1\) for a \(x \in E\), then \(P\) is said to be a sub-Markov kernel.

If for a fixed \(x \in E\), the \(P(x, \cdot)\) is a signed measure, then it is said to be a signed kernel. In that case, we will suppose that the signed kernel \(P\) is of bounded variation, that is, \(|P|(x, E) < \infty\).

If \(E\) is a finite or countable set, we take \(\mathcal{B}_E = \mathcal{P}_E\) (the set of all subsets of \(E\)), the Markov kernel is determined by the matrix \((P(i, j) : i, j \in E)\), with \(P(i, B) = \sum_{j \in B} P(i, j), B \in \mathcal{B}_E\).
2.1.1 Markov chains

A *time-homogeneous* Markov chain associated to a Markov kernel $P(x, B)$ is an adapted sequence of random variables $x_n$, $n \geq 0$ defined on $\Omega$, satisfying for every $n \in \mathbb{N}$, $x \in E$ and $B \in \mathcal{B}_E$, the following relation:

$$
P(x_{n+1} \in B | \mathcal{F}_n) = \mathbb{P}(x_{n+1} \in B | x_n) =: P(x_n, B), \quad (a.s.) \quad (2.1.1)$$

which is called the *Markov property*.

In most cases, we consider the Markov property with respect to $\mathcal{F}_n := \sigma(x_k, k \leq n)$, $n \geq 0$, the *natural filtration* generated by the chain $x_n$, $n \geq 0$.

If the Markov property (2.1.1) is satisfied for any finite $\mathcal{F}_n$-stopping time, then it is called *strong Markov property* and the chain is called a *strong Markov chain*.

The product of two Markov kernels $P$ and $Q$ defined on $(E, \mathcal{B}_E)$, is also a Markov kernel, defined by $PQ(x, B) = \int_E P(x, dy)Q(y, B)$.

Let us denote by $P^n(x, B) = \mathbb{P}(x_n \in B | x_0 = x) = \mathbb{P}(x_{n+m} \in B | x_m = x)$ the $n$-step transition probability which is defined inductively by the above relation. Using Markov property one gets the following property: for any $n, m \in \mathbb{N}$,

$$
P^{n+m}(x, B) = \int_E P^n(x, dy)P^m(y, B) = \int_E P^m(x, dy)P^n(y, B)$$

which is called the *Chapman-Kolmogorov equation*.

A subset $B \in \mathcal{B}_E$, is called *accessible* from a state $x \in E$, if $\mathbb{P}_x(x_n \in B$, for some $n \geq 1) > 0$ or equivalently $P^n(x, B) > 0$.

A Markov chain $x_n$, $n \geq 0$, is called *Harris recurrent* if there exists a $\sigma$-finite measure $\psi$ on $(E, \mathcal{B}_E)$, with $\psi(E) > 0$, such that for any $x \in E$, for any $A \in \mathcal{B}_E$ and $\psi(A) > 0$, $\mathbb{P}_x(\cup_{n \geq 1}\{x_n \in A\}) = 1$. 

If \( \mathbb{P}_x(\cup_{n \geq 1} \{ x_n \in A \}) \) is positive then the Markov chain is called \( \psi \)-irreducible.

The Markov chain is said to be uniformly irreducible if for any \( A \in \mathcal{B}_E \), \( \sup_x \mathbb{P}_x(\tau_A > N) \to \infty \) as \( N \to \infty \), where \( \tau_A := \inf \{ n \geq 0 : x_n \in A \} \).

A Markov chain \( x_n, n \geq 0 \), is said to be \( d \)-periodic \((d > 1)\), if there exists a cycle, that is, a sequence \((C_1, \ldots, C_d) \) of sets, \( C_i \in \mathcal{B}_E, 1 \leq i \leq d \), with \( P(x, C_{j+1}) = 1, x \in C_j, 1 \leq j \leq d - 1 \) and \( P(x, C_1) = 1, x \in C_d \), such that

- the set \( E \setminus \cup_{i=1}^d C_i \) is \( \psi \)-null
- if \((C'_1, \ldots, C'_d) \) is another cycle, then \( d' \) divides \( d \) and \( C'_i \) differs from a union of \( d/d' \) members of \((C_1, \ldots, C_d) \) only by a \( \psi \)-null set which is of type \( \cup_{i \geq 1} V_i \), where, for any \( i \geq 1 \), \( \mathbb{P}_x(\limsup \{ x_n \in V_i \}) = 0 \).

If \( d = 1 \) the the Markov chain is called aperiodic.

A probability measure \( \rho \) on \((E, \mathcal{B}_E)\) is said to be a stationary distribution or invariant probability for the Markov chain \( x_n, n \geq 0 \) if for any \( B \in \mathcal{B}_E \),

\[
\rho(B) = \int_E \rho(dx)P(x, B).
\]

If a Markov chain is \( \psi \)-irreducible and has invariant probability, it is called positive, otherwise it is called null.

If a Markov chain is Harris recurrent and positive it is called Harris positive.

If a Markov chain is aperiodic and Harris positive it is called (Harris) ergodic.

**Proposition 2.1** ([3]) Let \( x_n, n \geq 0, \) be an ergodic Markov chain, then

1) for any probability measure \( \alpha \) on \((E, \mathcal{B}_E)\), we have

\[
\| \alpha P^n - \rho \| \to 0, \quad n \to \infty
\]
2) for any $\varphi \in B(E)$,

$$\lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} \varphi(x_k) = \int_E \rho(dx) \varphi(x), \quad \mathbb{P}_\mu - a.s. \quad (2.1.2)$$

for any probability measure $\mu$ on $(E, \mathcal{B}_E)$.

Let us denote by $\Pi$ the stationary projector in $B(E)$ defined by the stationary distribution $\rho(B)$, $B \in \mathcal{B}_E$ of the Markov chain $x_n$ as:

$$\Pi \varphi(x) := \int_E \rho(dx) \varphi(y) \mathbf{1}(x) = \hat{\varphi}(x), \quad \hat{\varphi} := \int_E \rho(dx) \varphi(x),$$

where $\mathbf{1}(x) = 1$ for all $x \in E$. Then $\Pi^2 = \Pi$.

Let us denote by $P$ the operator of transition probabilities on $B(E)$ defined by

$$P \varphi(x) = \mathbb{E}[\varphi(x_{n+1})|x_n = x] = \int_E P(x,dy) \varphi(y)$$

and denote by $P^n$ the $n$-step transition operator corresponding to $P^n(x,B)$.

The Markov property (2.1.1) can be represented in the following form

$$\mathbb{E}[\varphi(x_{n+1})|\mathcal{F}_n] = P \varphi(x_n).$$

The Markov chain is uniformly ergodic if

$$\sup_{\|\varphi\| \leq 1} \|(P^n - \Pi) \varphi\| \to 0, \quad n \to \infty. \quad (2.1.3)$$

Uniform ergodicity of the Markov chain $x_n$ implies the convergence of the series

$$R_0 := \sum_{n=0}^{\infty} [P^n - \Pi] \quad (2.1.4)$$
which is called the potential operator of the Markov chain \( x_n, n \geq 0 \) (see [4]) and satisfies the property
\[
R_0[I - P] = [I - P]R_0 = I - \Pi.
\]

\[2.1.2\] Continuous-time Markov processes

Let us consider a family of Markov kernels \( \{P_t = P_t(x, B), t \in \mathbb{R}_+\} \). Let \( x(t) \geq 0 \) be an adapted stochastic process with values in \((E, \mathcal{B}_E)\) defined on \( \Omega \). The process \( x(t), t \geq 0 \), is said to be a time-homogeneous Markov process if for any fixed \( s, t \in \mathbb{R}_+ \) and \( B \in \mathcal{B}_E \),
\[
\mathbb{P}(x(t + s) \in B | \mathcal{F}_s) = \mathbb{P}(x(t + s) \in B | x(s)) = P_t(x(s), B), \quad (a.s.). \tag{2.1.5}
\]

When the Markov property (2.1.5) holds for any finite stopping time \( \tau \), instead of a deterministic time \( s \), we say that the Markov process \( x(t), t \geq 0 \), satisfies the strong Markov property and that the process \( x(t) \) is a strong Markov process. On the Banach space \( B(E) \), the operator of transition probability \( P_t \) is defined by
\[
P_t \varphi(x) = \mathbb{E}_x(\varphi(x(t))) = \int_E \varphi(y)P_t(x, dy), \; \varphi \in B(E)
\]
This is a contractive operator, i.e., \( \|P_t\varphi\| \leq \|\varphi\| \). The Chapman-Kolmogorov equation is equivalent to the following
\[
P_t P_s = P_{t+s}, \quad \text{for all } t, s \in \mathbb{R}_+
\]
which is called *semigroup property*. The Markov process \( x(t), t \geq 0 \) has a *stationary distribution* \( \pi \) if for any \( B \in \mathcal{B}_E \),

\[
\pi(B) = \int_E \pi(dx) P_t(x, B), \quad \pi(E) = 1, \quad t \geq 0.
\]

The Markov process \( x(t), t \geq 0 \) is *ergodic* if for every \( \varphi \in \mathcal{B}(E) \) we have

\[
\lim_{t \to \infty} \frac{1}{t} \int_0^t \varphi(x(s)) \, ds = \int_E \pi(dx) \varphi(x), \quad \mathbb{P}_\mu - a.s. \tag{2.1.6}
\]

for any probability \( \mu \) on \( (E, \mathcal{B}_E) \).

The *stationary projector* \( \Pi \), of an ergodic Markov process with stationary distribution \( \pi \), is defined as

\[
\Pi \varphi(x) := \int_E \pi(dx) \varphi(y) \mathbb{1}(x) = \hat{\varphi} \mathbb{1}(x), \quad \hat{\varphi} := \int_E \pi(dx) \varphi(x),
\]

where \( \mathbb{1}(x) = 1 \) for all \( x \in E \). We have \( \Pi^2 = \Pi \).

Let us consider a Markov process \( x(t), t \geq 0 \) with trajectories in \( \mathcal{D}[0, \infty) \) and semigroup \( \{P_t, t \geq 0\} \). The linear operator \( Q : \mathcal{B}(E) \to \mathcal{B}(E) \) defined by

\[
Q \varphi = \lim_{t \downarrow 0} \frac{1}{t} (P_t \varphi - \varphi),
\]

where the limit exists in norm, is called the *infinitesimal operator*. Let the domain of the operator \( \mathcal{D}(Q) \) be the subset of \( \mathcal{B}(E) \) for which the above limit exists. A Markov semigroup \( P_t, t \geq 0 \) is said to be *uniformly continuous* on \( \mathcal{B}(E) \), if

\[
\lim_{t \to 0} \|P_t - I\| = 0
\]
where \( I \) is the identity operator on \( B(E) \).

A time-homogeneous Markov process is said to be (purely) discontinuous or of jump type, if its semigroup is uniformly continuous. In that case, the process stays in any state for a positive (strict) time, and after leaving a state it moves directly to another one. This process we will call a jump Markov process.

Let \( x(t), t \geq 0 \) be a time-homogeneous jump Markov process. Let \( \tau_n, n \geq 0 \) be the jump times for which we have \( 0 = \tau_0 \leq \tau_1 \leq \cdots \leq \tau_n \leq \cdots \). The stochastic process \( x_n, n \geq 0 \) defined by \( x_n = x(\tau_n), n \geq 0 \) is called the embedded Markov chain of the Markov process \( x(t), t \geq 0 \).

Let \( P(x, B) \) be the transition probability of \( x_n, n \geq 0 \). The infinitesimal generator \( Q \) of the jump Markov process \( x(t), t \geq 0 \), is

\[
Q \varphi(x) = q(x) \int_E P(x, dy)[\varphi(y) - \varphi(x)]
\]

(2.1.7)

where the kernel \( P(x, B) \) is the transition kernel of the embedded Markov chain, and \( q(x), x \in E \) is the intensity of jumps function.

If the Markov process \( x(t) \) has a stationary distribution \( \pi \), then the embedded Markov chain \( x_n \) has also a stationary distribution \( \rho \) having the following property:

\[
\pi(dx)q(x) = q\rho(dx), \quad q := \int_E \rho(dx)q(x).
\]

**Example 2.2** The infinitesimal generator of a Poisson process.

Let \( S_1, S_2, \ldots \) be i.i.d. random variables such that \( \mathbb{P}(S_n > t) = e^{-\lambda t}, \lambda > 0 \). Let \( T_1 = S_1, T_n = S_1 + \cdots + S_n, n \geq 2 \). The process \( N_t, t \geq 0 \) defined by \( N_t = \sum_{n=1}^{\infty} \mathbb{I}(T_n \leq t) \) is a Markov process with respect to the natural filtration \( \mathcal{F}_t = \sigma\{N_s : s \leq t\} \) and its
trajectories $t \to N_t$ are right continuous step functions with jumps of height 1 at $T_n$ and they belong to the $\mathcal{D}(0, \infty)$. The random variable $N_t$ has Poisson distribution

$$\mathbb{P}(N_t = n) = e^{-\lambda t} \frac{(\lambda t)^n}{n!}, \quad \mathbb{E}(N_t) = \lambda t. \quad (2.1.8)$$

We can consider the time-homogeneous Poisson process $N_t$, $t \geq 0$ as a Markov family on the state $E = \mathbb{Z}_+ = \{0, 1, 2, \cdots \}$ with the probability measure $\mathbb{P}_x$ being such that $N_0 = x$, $x \in E$ and the transition probability function

$$\mathbb{P}(t, x, \{z\}) = \begin{cases} 
e^{-\lambda t} \frac{1}{(z-x)!}(\lambda t)^{z-x}, & z \geq x \\ 0, & z < x \end{cases}$$

Consider $\mathbb{L}$ to be the infinitesimal generator of this family. Then for any $\varphi \in B(E)$ we have

$$\mathbb{L}\varphi(x) = \lambda(\varphi(x + 1) - \varphi(x)).$$

**Example 2.3** *The infinitesimal generator of a renewal process.*

The renewal process is described similarly as Poisson process but the interarrival times $S_n$ are now random variables with a distribution function $f$ on $\mathbb{R}_+$ not necessarily exponential. The times $T_n$ are called renewal times. The function $m(t) = \mathbb{E}(N_t)$, the expected value of the counting process at time $t$, is called the renewal function and it verifies the following renewal equation:

$$m(t) = F(t) + \int_0^t F(t - s) \, dm(s), \quad F(t) = \int_0^t f(s) \, ds$$
The renewal process is not a Markov process, but if we denote the last time jump by \( \nu(t) = \sup\{k : T_k \leq t\} \), and the time since the last jump by \( \tau_t = t - T_{\nu(t)} \), then the coupled process \( x_t = (N_t, \tau_t) \), \( t \geq 0 \) becomes a Markov process in \( \mathbb{Z}_+ \times \mathbb{R}_+ \).

The hazard rate corresponding to the renewal density \( f \) is \( \lambda(t) = \frac{f(t)}{\bar{F}(t)} \), where \( \bar{F}(t) = \int_t^\infty f(s)ds \) is the survivor function. If the start point is \( x = (n, \tau) \) then after a short time \( \delta \) the process is at \( (n, \tau + \delta) \) or one jump to \( (n + 1, \delta') \) for some \( \delta' \in (0, \delta) \). Therefore, for any bounded measurable function \( \varphi : E \to \mathbb{R} \) we find, under the assumption that \( \tau \to \varphi(n, \tau) \), that the infinitesimal generator of the process \( x(t), t \geq 0 \) has the following form

\[
\mathbb{L}\varphi(n, \tau) = \lim_{\delta \to 0} \frac{\mathbb{E}(n, \tau)\varphi(x_\delta) - \varphi(n, \tau)}{\delta} = \frac{d}{d\tau}\varphi(n, \tau) + \lambda(\tau)[\varphi(n + 1, 0) - \varphi(n, \tau)]
\]

**Example 2.4** Coupled Markov process

Let \( x(t), t \geq 0 \) be a nonhomogeneous jump Markov process. Then the infinitesimal generator of the coupled Markov process \( (t, x(t)), t \geq 0 \) is defined as follows:

\[
\mathbb{L}\varphi(t, x) = \frac{\partial}{\partial t}\varphi(t, x) + Q\varphi(\cdot, x)
\]

(2.1.9)

where \( Q \) is defined in (13) and the domain of \( \mathbb{L} \) is \( \mathcal{D}(\mathbb{L}) = C^{1,0}(\mathbb{R}_+ \times E) \)

**Example 2.5** Increment process

Let \( x(t), t \geq 0 \) be a pure jump Markov process (that is, without drift and diffusion part) with state space \( E \) and infinitesimal generator \( Q \), and let \( \nu(t), t \geq 0 \) be the corresponding counting process of jumps and \( x_n, n \geq 0 \) be the embedded Markov chain. Let \( a \) be a real-valued measurable function on the state space \( E \) and consider the increment process

\[
\alpha(t) = \sum_{k=1}^{\nu(t)} a(x_k), \ t \geq 0
\]
Then the infinitesimal generator of the coupled Markov process \((\alpha(t), x(t)), t \geq 0\) is

\[
\mathbb{L} = Q + Q_0[I \Gamma(x) - I]
\]

where \(\Gamma(x)\varphi(u) := \varphi(u + a(x))\), \(Q_0\varphi(x) := q(x) \int_E P(x, dy)\varphi(y)\) and \(I\) is the identity operator.

**Example 2.6** *Integral functionals [5]*

Let the jump Markov process \(x(t), t \geq 0\) have the infinitesimal generator \(Q\) as in previous example and let us consider the process \(\xi(t) := \int_0^t a(x(s)) \, ds\). Then the infinitesimal generator of the coupled process \((\xi(t), x(t)), t \geq 0\) is \(\mathbb{L} = Q + A(x)\), where \(A(x)\varphi(u) := a(x)\varphi'(u)\).

### 2.1.3 Diffusion processes

A nonhomogeneous Markov process \(x(t), t \geq 0\) defined on \(\Omega\) with values in \(\mathbb{R}^d, d \geq 1\), with transition function \(P_{s,t}(x, B) := \mathbb{P}(x(t) \in B| x(s) = x)\) for \(0 \leq s < t < \infty, x \in \mathbb{R}^d, B \in \mathcal{B}_d\), is said to be a *diffusion process*, with infinitesimal generator \(\mathbb{L}_t\) if

- (a) it has continuous paths
- (b) for any \(x \in \mathbb{R}^d\) and any \(\varphi \in C^2(\mathbb{R}^d)\),

\[
\int_{\mathbb{R}^d} P_{t,t+h}(x, dy)[\varphi(y) - \varphi(x)] = h\mathbb{L}_t\varphi(x) + o(h), \ h \to 0.
\]

Let \(a\) be a real-valued measurable function, defined on \(\mathbb{R}^d \times \mathbb{R}_+\), and \(B\) a function defined on \(\mathbb{R}^d \times \mathbb{R}_+\) with values in the space of symmetric positive operators from \(\mathbb{R}^d\) to \(\mathbb{R}^d\).
The infinitesimal generator $\mathbb{L}_t$ of the diffusion $x(t), t \geq 0$ acts on functions $\varphi$ in $C^2(\mathbb{R}^d)$ as follows

$$\mathbb{L}_t \varphi(x) = a(x,t) \varphi'(x) + \frac{1}{2} B(x,t) \varphi''(x).$$ (2.1.10)

The drift coefficient $a(x,t)$ and the diffusion operator (or coefficient) $B(x,t)$ apply as follows:

$$a(x,t) \varphi'(x) = \sum_{i=1}^{d} a_i(x,t) \frac{\partial}{\partial x_i} \varphi(x)$$

and

$$B(x,t) \varphi''(x) = \sum_{i,j=1}^{d} b_{ij}(x,t) \frac{\partial^2}{\partial x_i \partial x_j} \varphi(x).$$

For a time-homogeneous diffusion the infinitesimal generator does not depend on $t$, that means that functions $a$ and $B$ are free of $t$. If $a(x,t) \equiv 0$ and $B(x,t) = I$ then $x(t), t \geq 0$ is a Wiener process.

2.1.4 Processes with independent increments

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ be a stochastic basis, and $x(t), t \geq 0$ be a stochastic process adapted to the filtration $(\mathcal{F}_t)_{t \geq 0}$ with values in $\mathbb{R}^d$.

The process $x(t), t \geq 0$, is said to be with independent increments if for any $s, t \in \mathbb{R}_+$ with $s < t$, $x(t) - x(s)$ is independent of $\mathcal{F}_s$.

If the filtration $(\mathcal{F}_t)_{t \geq 0}$ is the natural filtration of the process then this property is equivalent to the increment independence property: for any $n \geq 1$ and any $0 \leq t_0 < t_1 < \cdots < t_n < \infty$, the random variables $x(t_0), x(t_1) - x(t_0), \ldots, x(t_n) - x(t_{n-1})$ are independent.
If moreover, \( x(t) \) has stationary increments, which means, the law of \( x(t) - x(s) \) depends only on \( t - s \), then the process \( x(t) \) is said to be a process with stationary independent increments.

The main properties of the processes with stationary independent increments are that their distributions are infinitely divisible and have the Markov property. Let

\[
\phi_t(\lambda) := \mathbb{E} \exp(i\lambda(x(t + s) - x(s)))
\]

be the characteristic function of the increments. Then it satisfies the semigroup property

\[
\phi_{t+s}(\lambda) = \phi_t(\lambda)\phi_s(\lambda)
\]

and can be represented as \( \phi_t(\lambda) = \exp[t\psi(\lambda)] \). Then the cumulant \( \psi(\lambda) \) has the well known Lévy-Khintchine formula ([3]):

\[
\psi(\lambda) = i\lambda a - \frac{1}{2}\sigma^2\lambda^2 + \int_\mathbb{R} [e^{i\lambda z} - 1 - i\lambda z\mathbb{1}_{\{|z| \leq 1\}}] H(dz),
\]

(2.1.11)

where \( H \) is the spectral measure and satisfies the following conditions

\[
\int_{|z| \leq 1} z^2 H(dz) < \infty, \quad \int_{|z| > 1} H(dz) < \infty
\]

The processes with stationary independent increments satisfy the Markov property, that is, the transition probabilities are generated by the Markov semigroup

\[
\Gamma_t \varphi(u) := \mathbb{E}\varphi(u + x(t)).
\]

(2.1.12)
The generator of the semigroup (2.1.12) with the cumulant (2.1.11) has the following representation

\[ \Gamma \varphi(u) = a \varphi'(u) - \frac{\sigma^2}{2} \varphi''(u) + \int_{\mathbb{R}} \left[ \varphi(u + v) - \varphi(u) - v \varphi'(u) \mathbb{1}_{|v|\leq 1} \right] H(dv). \]

The most important processes with stationary independent increments are Brownian motion, Poisson process, and Lévy process.

The *standard Wiener process* with the cumulant \( \psi(\lambda) = -\sigma^2 \lambda^2 / 2 \) has the generator \( \Gamma \varphi(u) = \sigma^2 \varphi''(u) / 2 \).

The *Compound Poisson process* \( x(t) = \sum_{\nu(t)} \xi_k \), where \( \nu(t), t \geq 0 \) is a homogeneous Poisson process with intensity \( \lambda > 0 \) and \( \xi_k, k \geq 1 \) is an i.i.d. sequence of real random variables, independent of \( \nu(t), t \geq 0 \), with common distribution function \( F \), has the cumulant of the form

\[ \psi(\lambda) = \lambda \int_{\mathbb{R}} [e^{i\lambda z} - 1] F(dz) \]

and the infinitesimal generator of the form

\[ \Gamma \varphi(u) = \lambda \int_{\mathbb{R}} [\varphi(u + v) - \varphi(u)] F(dv). \]

One can associate uniquely a homogeneous Markov process to the process with independent increments \( x(t) \) by

\[ \xi(t) := x(s + t) - x(s) + u \quad s \geq 0, u \in \mathbb{R}^d. \]
The semigroup $\Gamma_t \varphi(u) = \mathbb{E}_u \varphi(\xi(t)) = \mathbb{E}_u \varphi(u + x(t))$ has the following remarkable property: let $S_v : \mathcal{B}_d \rightarrow \mathcal{B}_d$ be the shift operator, $S_v \varphi(u) = \varphi(u + v)$; then $\Gamma_t S_v \varphi(u) = S_v \Gamma_t \varphi(u)$ for any $u, v \in \mathbb{R}^d$.

The Markov processes with the property that their semigroup commutes with the operator $S_v$, $v \in \mathbb{R}^d$ are called homogeneous in space.

Thus, processes with stationary independent increments can be identified in a certain sense with Markov processes that are homogeneous in both time and space.

### 2.1.5 Processes with locally independent increments

The processes with locally independent increments considered are jump Markov processes with drift and without diffusion part. These processes are also called weakly differentiable Markov processes [3] or piecewise-deterministic Markov processes [6]. It is worth noticing that such processes include strictly the independent increment processes. For their detailed presentation and applications see [6].

These processes are defined by the infinitesimal generator $\Pi$ as follows

$$\Pi \varphi(u) = a(u) \varphi'(u) + \int_{\mathbb{R}^d} [\varphi(u + v) - \varphi(u) - v \varphi'(u)] \Gamma(u, dv)$$

(2.1.13)

with drift velocity $a(u)$, the intensity kernel $\Gamma(u, dv)$ such that $\Gamma(u, dv) \in \mathbb{R}_+$ and satisfying the above conditions of the spectral measure. It is understood that when $d > 1$, we have

$$v \varphi'(u) = \sum_{k=1}^d v_k \frac{\partial \varphi}{\partial u_k}(u).$$

Let us be given the Euclidian space $\mathbb{R}^d$ with the Borel $\sigma$-algebra $\mathcal{B}_d$ and the compact measurable space $(E, \mathcal{E})$. We consider the family of time-homogeneous Markov processes $\eta(t; x)$, $t \geq 0$, $x \in E$, with trajectories in $\mathcal{D}[0, \infty)$, with locally independent
increments. These processes take values in the Euclidian space \( \mathbb{R}^d \) \( (d \geq 1) \), and depend on the state \( x \in E \), and their infinitesimal generators are given by

\[
\Gamma(x)\varphi(u) = a(u; x)\varphi'(u) + \int_{\mathbb{R}^d} [\varphi(u + v) - \varphi(u) - v\varphi'(u)]\Gamma(u, dv; x) \tag{2.1.14}
\]

These processes will be used in Chapter 3 as \textit{switched processes}. The drift velocity \( a(u; x) \) and the measure of the random jumps \( \Gamma(u, dv; x) \) depend on the state \( x \in E \).

2.2 Weak convergence of stochastic processes

2.2.1 Martingale characterization of Markov processes

Let \((\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})\) be a probability space and let \( M_t, t \geq 0 \) be a process on \( \Omega \) and adapted to the filtration \((\mathcal{F}_t)_{t \geq 0}\). A real-valued process \( M_t, t \geq 0 \) adapted to the filtration \((\mathcal{F}_t)_{t \geq 0}\) is called a \textit{martingale} (\textit{submartingale}, \textit{supermartingale}) if

(i) \( \mathbb{E}|M_t| < \infty \) for all \( t \geq 0 \)

(ii) for any \( s < t \),

\[
\mathbb{E}(M_t|\mathcal{F}_s) = M_s, \ (\mathbb{E}(M_t|\mathcal{F}_s) \geq M_s, \ \mathbb{E}(M_t|\mathcal{F}_s) \leq M_s), \ a.s.
\]

The process \( M_t, t \geq 0 \) is called a \textit{square integrable martingale} if

\[
\sup_{t \geq 0} \mathbb{E}M_t^2 < \infty
\]

If \( M_t, t \geq 0 \) is a square integrable martingale then the process \( M_t^2, t \geq 0 \) is an \( \mathcal{F}_t \)-submartingale, and the \textit{Doob-Meyer decomposition} \[7\] holds

\[
M_t^2 = \langle M_t \rangle + N_t
\]
where \( N_t, t \geq 0 \) is a martingale and the increasing process \( \langle M_t \rangle, t \geq 0 \) is called the \textit{square characteristic} of the martingale \( M_t \).

A process \( M_t, t \geq 0 \) is called \textit{local martingale} if there exists an increasing sequence of \textit{stopping times} \( \tau_n \) such that the stopped process \( M^\tau_n := M_{t \wedge \tau_n}, t \geq 0 \) is a martingale for each \( n \geq 1 \).

Let \( x(t), t \geq 0 \) be a Markov process with a standard state space \((E, \mathcal{E})\), defined on \( \Omega \). Let \( P_t(x, B), x \in E, B \in \mathcal{E}, t \geq 0 \) be its transition function, and \( P_t, t \geq 0 \), its strongly continuous semigroup defined on the Banach space \( B(E) \) of real-valued measurable functions defined on \( E \). Let \( Q \) be the infinitesimal generator of the semigroup \( P_t, t \geq 0 \), with the dense domain of definition \( \mathcal{D}(Q) \subset B \).

For any function \( \varphi \in \mathcal{D}(Q) \subset B(E) \) and \( t > 0 \) we have the \textit{Dynkin formula}

\[
P_t\varphi(x) = \varphi(x) + \int_0^t QP_s\varphi(x) \, ds.
\]

From this formula, using conditional expectation, we get

\[
\mathbb{E}_x[\varphi(x(t)) - \varphi(x) - \int_0^t Q\varphi(x(s)) \, ds] = 0.
\]

Thus, the process

\[
\mu(t) := \varphi(x(t)) - \varphi(x) - \int_0^t Q\varphi(x(s)) \, ds \tag{2.2.1}
\]

is an \( \mathcal{F}^x_t = \sigma(x(s), s \leq t) \)-martingale.

The following theorems gives the \textit{martingale characterization} of Markov processes.

**Theorem 2.7 \([8]\)** Let \((E, \mathcal{E})\) be a standard state space and let \( x(t), t \geq 0 \) be a stochastic process on it, adapted to the filtration \( (\mathcal{F}_t)_{t \geq 0} \). Let \( Q \) be the generator of
a strongly continuous semigroup $P_t$, $t \geq 0$, on the Banach space $B(E)$, with dense domain $D(Q) \in B(E)$. If for any $\varphi \in D(Q)$, the process $\mu(t)$, $t \geq 0$, defined by (2.2.1) is an $\mathcal{F}_t$-martingale, then $x(t)$, $t \geq 0$ is a Markov process generated by the infinitesimal generator $Q$.

The process $x(t)$, $t \geq 0$ is said to solve the martingale problem for the infinitesimal generator $Q$.

Let $x_n$, $n \geq 0$ be a Markov chain on a measurable state space $(E,\mathcal{E})$ induced by a stochastic kernel $P(x,B)$, $x \in E$, $B \in \mathcal{E}$. Let $P$ be the corresponding transition operator defined on the Banach space $B(E)$.

Let us construct now the following martingale as a sum of martingale differences

$$
\mu_n = \sum_{k=0}^{n-1} [\varphi(x_{k+1}) - \mathbb{E}(\varphi(x_{k+1})|\mathcal{F}_k)].
$$

By using the Markov property and the rearrangement of terms the martingale is written as

$$
\mu_n = \varphi(x_n) - \varphi(x_0) - \sum_{k=1}^{n-1} [P-I] \varphi(x_k). \tag{2.2.2}
$$

This representation of the martingale is associated with a Markov chain characterization.

**Proposition 2.8** Let $x_n$, $n \geq 0$ be a sequence of random variables taking values in a measurable space $(E,\mathcal{E})$ and adapted to the filtration $\mathcal{F}_n$, $n \geq 0$. Let $P$ be a bounded linear positive operator on the Banach space $B(E)$ induced by a transition probability kernel $P(x,B)$ on $(E,\mathcal{E})$. If for every $\varphi \in B(E)$, the right hand side of (2.2.2) is a martingale then the sequence $x_n$, $n \geq 0$ is a Markov chain with transition probability kernel $P(x,B)$ induced by the operator $P$. 
2.2.2 Reducible-invertible operators

A bounded linear operator $Q : B(E) \to B(E)$ is called reducible-invertible if the Banach space $B(E)$ can be decomposed in direct sum of two subspaces:

$$B(E) = N_Q \oplus R_Q, \quad \dim N_Q \geq 1$$  \hspace{1cm} (2.2.3)

where $N_Q := \{ \varphi : Q\varphi = 0 \}$ is the null space, and $R_Q := \{ \psi : Q\varphi = \psi, \varphi \in B(E) \}$ is the range space. [4]

The decomposition (2.2.3) generates the projector $\Pi$ on $N_Q$

$$\Pi \varphi := \begin{cases} 
\varphi, & \varphi \in N_Q \\
0, & \varphi \in R_Q
\end{cases}$$

The operator $I - \Pi$ is a projector on the subspace $R_Q$

$$(I - \Pi) \varphi := \begin{cases} 
0, & \varphi \in N_Q \\
\varphi, & \varphi \in R_Q
\end{cases}$$

where $I$ is the identity operator in $B(E)$.

Let $Q$ be a reducible-invertible operator. The operator

$$R_0 := (Q + \Pi)^{-1} - \Pi$$  \hspace{1cm} (2.2.4)

is called the potential operator of $Q$. The potential can also be written as $R_0 = (\Pi - Q)^{-1} - \Pi$. 


Proposition 2.9 ([4]) The following equalities hold:

\[ QR_0 = R_0Q = I - \Pi \]
\[ \Pi R_0 = R_0\Pi = 0 \]
\[ QR_0^n = R_0^nQ = R_0^{n-1}, \quad n \geq 1. \]

Proposition 2.10 ([4]) Let \( Q : B(E) \to B(E) \) be a reducible-invertible operator.

The equation

\[ Q\varphi = \psi \quad (2.2.5) \]

under the solvability condition \( \Pi \psi = 0 \) has a general solution of the form:

\[ \varphi = -R_0\psi + \varphi_0, \quad \varphi_0 \in N_Q. \quad (2.2.6) \]

If moreover, the condition \( \Pi \varphi = 0 \) holds then the equation has a unique solution represented by

\[ \varphi = -R_0\psi. \quad (2.2.7) \]

Note that for a uniformly ergodic Markov chain the operator \( Q := P - I \) is reducible-invertible. For a uniformly ergodic Markov process with infinitesimal generator \( Q \) and semigroup \( P_t, \quad t \geq 0 \) the potential \( R_0 \) is a bounded operator defined by

\[ R_0 = \int_0^\infty (P_t - \Pi)dt, \quad (2.2.8) \]

where the projector \( \Pi \) is defined as

\[ \Pi \varphi(u) = \int_E \rho(dx)\varphi(x)\mathbb{I}(x), \]
with \( \rho(B) \), \( B \in \mathcal{E} \) the stationary distribution of the Markov chain.

Let \( Q_\varepsilon : B(E) \rightarrow B(E), \varepsilon > 0 \), be a family of linear operators. We say that it converges in the strong sense as \( \varepsilon \rightarrow 0 \) to the operator \( Q \) if

\[
\lim_{\varepsilon \rightarrow 0} \|Q_\varepsilon \varphi - Q \varphi\| = 0
\]

for any \( \varphi \in B(E) \). We use the notation \( s - \lim_{\varepsilon \rightarrow 0} Q_\varepsilon = Q \).

### 2.2.3 Perturbation of reducible-invertible operators

Let \( Q \) be a bounded reducible-invertible operator on \( B(E) \).

The problem of asymptotic singular perturbation of a reducible-invertible operator ([4]) \( Q \) with small parameter \( \varepsilon > 0 \) and perturbing operator \( Q_1 \) is formulated in the following way: we have to construct the vector \( \varphi^\varepsilon = \varphi + \varepsilon \varphi_1 \) which realizes the asymptotic representation

\[
[\varepsilon^{-1}Q + Q_1] \varphi^\varepsilon = \psi + \varepsilon \theta_\varepsilon
\]

for some given vector \( \psi \) and uniformly bounded in norm vector \( \theta_\varepsilon \), \( \|\theta_\varepsilon\| \leq c \) as \( \varepsilon \rightarrow 0 \).

**Proposition 2.11** ([4]) Let the following conditions be satisfied:

(a) the bounded operator \( Q \) is reducible-invertible

(b) the perturbing operator \( Q_1 \) is closed with a dense domain \( B_0(E) \subseteq B(E) \)

(C) the contracted operator \( \hat{Q}_1 \), defined by \( \Pi Q_1 \Pi = \hat{Q}_1 \Pi \), has the inverse operator \( \hat{Q}_1^{-1} \).
Then the asymptotic representation (2.2.9) is realized by the vectors which are determined by

\[ \hat{Q}_1\hat{\varphi} = \hat{\psi} \]
\[ \varphi_1 = R_0(\psi - Q_1\varphi) \]
\[ \theta_\varepsilon = Q_1R_0(\psi - Q_1\varphi) \]
\[ R_0 = (Q + \Pi)^{-1} - \Pi \]

Proof: The left hand side of (2.2.9) can be represented in the form:

\[ [\varepsilon^{-1}Q + Q_1](\varphi + \varepsilon\varphi_1) = \varepsilon^{-1}Q\varphi + (Q\varphi_1 + Q_1\varphi) + \varepsilon Q_1\varphi_1 \]

In order to realize the right-hand side of (2.2.9) we set

\[ Q\varphi = 0 \]
\[ Q\varphi_1 + Q_1\varphi = \psi \]
\[ Q_1\varphi_1 = \theta_\varepsilon \]

The first relation implies that \( \varphi \in N_Q \), from the third we get that \( \theta_\varepsilon = Q_1\varphi_1 \) is independent of \( \varepsilon \). The main problem is to solve the equation

\[ Q\varphi_1 = \psi - Q_1\varphi \quad (2.2.10) \]

The solvability of (2.2.10) is \( \Pi(\psi - Q_1\varphi) = 0 \). Since \( \Pi\varphi = \varphi \) we have

\[ \Pi\psi = \Pi Q_1\Pi\varphi \quad (2.2.11) \]
On the contracted space $\hat{N}_Q$, introduce the contracted operator $\hat{Q}_1$ which is determined by relation $\Pi Q_1 \Pi = \hat{Q}_1 \Pi$ and set $\hat{\psi} := \hat{\Pi} \psi \in \hat{N}_Q$. Thus, the relation $\hat{Q}_1 \hat{\varphi} = \hat{\psi}$ establishes a connection between two vectors $\hat{\varphi}$ and then $\hat{\psi}$ in $\hat{N}_Q$.

The solution of (2.2.10) is $\varphi_1 = R_0(\psi - Q_1 \varphi)$ where $R_0$ is the potential operator of $Q$, and $\theta_\varepsilon = Q_1 \varphi_1 = Q_1 R_0(\psi - Q_1 \varphi)$.

### 2.2.4 Weak convergence

Consider the functional space $D_E[0, \infty)$ as the space of sample paths of stochastic processes.

The main type of convergence of stochastic processes is the weak convergence of finite-dimensional distributions, that is for the family of stochastic processes $x^\varepsilon(t)$, $t \geq 0$, $\varepsilon > 0$ there is a process $x(t)$, $t \geq 0$ such that

$$\lim_{\varepsilon \to 0} \mathbb{E} \varphi(x^\varepsilon(t_1), \ldots, x^\varepsilon(t_N)) = \mathbb{E} \varphi(x(t_1), \ldots, x(t_N))$$

for any test function $\varphi(x_1, \ldots, x_N) \in C_b(E^N)$, the space of real-valued bounded continuous functions on $E^N$, and for any finite set $\{t_1, \ldots, t_N\} \in S$, where $S$ is a dense set in $\mathbb{R}_+$. We will denote this convergence by

$$x^\varepsilon(t) \overset{D}{\to} x(t), \quad \varepsilon \to 0.$$

A more general type of convergence of stochastic processes is the weak convergence of associated measures, that is, the probability distributions

$$P_\varepsilon(B) := \mathbb{P}(x^\varepsilon(\cdot) \in B), \quad B \in D_E$$
where $D_E$ is the Borel $\sigma$-algebra on $D_E[0, \infty)$ converge weakly to $P(B)$. Therefore, the following limit holds

$$\lim_{\varepsilon \to 0} \mathbb{E}\varphi(x^\varepsilon(\cdot)) = \mathbb{E}\varphi(x(\cdot))$$

for all $\varphi \in C_b(D_E[0, \infty))$, the space of all bounded continuous real-valued functions on $D_E[0, \infty)$. We will denote the weak convergence of processes and that of the associated measures by

$$x^\varepsilon \Rightarrow x, \quad P^\varepsilon \Rightarrow P, \quad \varepsilon \to 0.$$

Theorem 2.12 ([9]) Let $x^\varepsilon(t), t \geq 0, \varepsilon > 0$, and $x(t), t \geq 0$, be processes with sample paths in $D_E[0, \infty)$.

(a) if $x^\varepsilon \Rightarrow x$ as $\varepsilon \to 0$, then $x^\varepsilon \overset{D}{\Rightarrow} x$, for the set $S = \{t > 0 : \mathbb{P}(x(t) = x(t_-)) = 1\}$.

(b) if $(x^\varepsilon, \varepsilon > 0)$ is relatively compact and there exists a dense set $S \subset \mathbb{R}_+$ such that $x^\varepsilon \overset{D}{\Rightarrow} x$ as $\varepsilon \to 0$ on the set $S$, then $x^\varepsilon \Rightarrow x$ as $\varepsilon \to 0$.

Theorem 2.13 (Prohorov’s theorem [10]) The relative compactness of a family of probability measures on $D_E[0, \infty)$ is equivalent to the tightness of this family.

For Markov processes, the best way to verify the relative compactness is to use the martingale characterization approach of Stroock and Varadhan, described in the following theorem.

Theorem 2.14 [8] Let the family of $\mathbb{R}^d$-valued stochastic processes $x^\varepsilon(t), t \geq 0, \varepsilon > 0$ such that:
H1: for all nonnegative functions $\varphi \in C_0^\infty(\mathbb{R}^d)$ there exists a constant $A_\varphi \geq 0$ such that $(\varphi(x^\varepsilon(t)) + A_\varphi(t), \mathcal{F}_t^\varepsilon)$ is a nonnegative submartingale.

H2: given a nonnegative $\varphi \in C_0^\infty(\mathbb{R}^d)$, the constant $A_\varphi$ can be chosen such that it does not depend on the translates of $\varphi$.

Then, under the initial condition

$$\lim_{c \to \infty} \sup_{0 < \varepsilon \leq \varepsilon_0} \mathbb{P}(\|x^\varepsilon(0)\| \geq c) = 0,$$

the family of associated probability measures $P_\varepsilon$, $\varepsilon > 0$ is relatively compact.

It is worth noticing that the relative compactness conditions formulated in $\mathcal{C}_0[0, \infty)$ are valid for the space $\mathcal{D}_0[0, \infty)$.

In order to verify the weak convergence of a family of stochastic processes in $\mathcal{D}_E[0, \infty)$ we have to establish the relative compactness and the weak convergence of finite-dimensional distributions. For a family of Markov processes these are derived from the martingale characterization of Markov processes and the convergence of their infinitesimal generators.

**Theorem 2.15 ([5])** Let $x^\varepsilon(t), t \geq 0, \varepsilon > 0$ be a family of stochastic processes with values in the standard phase space $(E, \mathcal{E})$ and adapted to the filtration of $\sigma$-algebras $\mathcal{F}_t^\varepsilon, t \geq 0, \varepsilon > 0$. Let $x(t)$ be a homogeneous stochastically continuous Markov process whose transition probabilities are uniquely determined by the generating operator $L$ which is defined on a dense domain $\mathcal{D}_L$ in the space $\mathcal{C}_0(E)$ and in addition assume that $\mathcal{C}_0(E)$ is contained in the closure of $\mathcal{D}_L$. Suppose that for all $\varphi \in \mathcal{D}_L$

$$\varphi(x^\varepsilon(t)) - \int_0^t L\varphi(x^\varepsilon(s)) \, ds = \mu_t^\varepsilon + \psi_t^\varepsilon$$

(2.2.12)
where \((\mu_t^\varepsilon, \mathcal{F}_t^\varepsilon, t \geq 0)\) is a family of square integrable martingales with the quadratic characteristic admitting representation in the form

\[
\langle \mu_t^\varepsilon \rangle_t = \int_0^t \zeta^\varepsilon(s) \, ds
\]

with random functions \(\zeta^\varepsilon(s), s \geq 0\) satisfying

\[
\mathbb{E} \sup_{0 \leq t \leq T} |\zeta^\varepsilon(s)|^{1+\delta} \leq c < \infty \tag{2.2.13}
\]

and the process \(\psi_t^\varepsilon, \varepsilon > 0\) verifying

\[
\sup_{\varepsilon \leq \varepsilon_0} \mathbb{E} \sup_{0 \leq t \leq T} |\psi_t^\varepsilon| \to 0 \quad \text{as} \quad \varepsilon \to 0
\]

Then \(x^\varepsilon \to x\) as \(\varepsilon \to 0\) and the limiting process \(x(t), t \geq 0\) is characterized by the martingale

\[
\varphi(x(t)) - \int_0^t L\varphi(x(s)) \, ds = \mu_0^t.
\]

If the condition (2.2.13) is replaced by

\[
\mathbb{E} \sup_{0 \leq t \leq T} |\zeta^\varepsilon(s)| \to 0 \quad \text{as} \quad \varepsilon \to 0
\]

then the limiting process \(x(t)\) is given by the solution of the deterministic evolutionary equation

\[
\frac{d\varphi(x(t))}{dt} = L\varphi(x(t)).
\]

### 2.3 Large deviation principle

Let \(\{X^n, n \in \mathbb{N}\}\) be a sequence of random variables defined on a probability space \((\Omega, \mathcal{F}, \mathbb{P})\) and taking values in a complete, separable metric space \((\mathcal{X}, d)\). The
theory of large deviations focuses on random variables \( \{X^n\} \) for which the probabilities \( \mathbb{P}\{X^n \in A\} \) converge to 0 exponentially fast for a class of Borel sets \( A \). The exponential decay rate of these probabilities is expressed in terms of a function \( I \) mapping \( \mathcal{X} \) into \([0, \infty]\). This function is called a rate function if it has compact level sets, i.e., for each \( M < \infty \) the level set \( \{x \in \mathcal{X} : I(x) \leq M\} \) is a compact subset of \( \mathcal{X} \). A function having compact level sets is automatically lower semicontinuous and it attains its infimum on any nonempty closed set.

**Definition 2.16** The sequence \( \{X^n\} \) is said to satisfy the large deviation principle on \( \mathcal{X} \) with rate function \( I \) if the following two conditions hold:

(a) Large deviation upper bound: for each closed subset \( F \) of \( \mathcal{X} \)

\[
\limsup_{n \to \infty} \frac{1}{n} \log \mathbb{P}\{X^n \in F\} \leq - \inf_{x \in F} I(x) \]

(b) Large deviation lower bound: for each open subset \( G \) of \( \mathcal{X} \)

\[
\liminf_{n \to \infty} \frac{1}{n} \log \mathbb{P}\{X^n \in G\} \geq - \inf_{x \in G} I(x). \]

**Definition 2.17** Let \( I \) be a rate function on \( \mathcal{X} \). The sequence \( \{X^n\} \) is said to satisfy the Laplace principle on \( \mathcal{X} \) with rate function \( I \) if for all bounded continuous functions \( h \)

\[
\lim_{n \to \infty} \frac{1}{n} \log \mathbb{E}\{\exp[-nh(X^n)]\} = - \inf_{x \in \mathcal{X}} \{h(x) + I(x)\}. \quad (2.3.1)
\]

The large deviation principle can be formulated in terms of the Laplace principle by using the following fundamental results.

**Theorem 2.18** (Varadhan’ theorem [11]) Assume that the sequence \( \{X^n\} \) satisfies the large deviation principle on \( \mathcal{X} \) with rate function \( I \). Then for all bounded con-
tinuous functions $h$ mapping $\mathcal{X}$ into $\mathbb{R}$, the equation (2.3.1) holds and therefore the sequence $\{X^n\}$ satisfies the Laplace principle on $\mathcal{X}$.

A converse of Varadhan’s theorem is also due to Varadhan and it states:

**Theorem 2.19** ([11]) The Laplace principle implies the large deviation principle with the same rate function. More precisely, if $I$ is a rate function on $\mathcal{X}$ and the limit 2.3.1 is valid for all bounded continuous functions $h$, then $\{X^n\}$ satisfies the large deviation principle on $\mathcal{X}$ with rate function $I$.

**Definition 2.20** The sequence $\{X^n\}$ is said to be exponentially tight if for each $M \in (0, \infty)$ there exists a compact set $K$ such that

$$
\limsup_{n \to \infty} \frac{1}{n} \log \mathbb{P}\{X^n \in K^c\} \leq -M.
$$

The next theorem is another converse to Varadhan’s theorem due to Bryc.

**Theorem 2.21** (Bryc’s theorem [12]) If the sequence $\{X^n\}$ is exponentially tight and if the limit

$$
\Lambda(h) := \lim_{n \to \infty} \log \mathbb{E}\{\exp[-nh(X^n)]\}
$$

exists for all $h$ bounded continuous functions, then $\{X^n\}$ satisfies the large deviation principle on $\mathcal{X}$ with rate function

$$
I(x) := -\inf_{h \in C_b(\mathcal{X})} \{h(x) + \Lambda(h)\}
$$

and $\Lambda(h) = -\inf_{x \in \mathcal{X}} \{h(x) + I(x)\}$.

If a sequence of random variables satisfies the Laplace principle (or equivalently the LDP) with some rate function, then the rate function is unique.

**Theorem 2.22** ([11]) Assume that $\{X^n\}$ satisfies the Laplace principle on $\mathcal{X}$ with rate function $I$ and with rate function $J$. Then $I(\xi) = J(\xi)$ for all $\xi \in \mathcal{X}$.
The continuous image of a sequence of random variables satisfying the Laplace principle (or LDP) also satisfies the Laplace principle (or LDP). This property is illustrated in the next theorem.

**Theorem 2.23** (Contraction Principle [11]) Let \( X \) and \( Y \) be Polish spaces, \( f : X \to Y \) a continuous function and \( I \) a rate function on \( X \).

(a) For each \( y \in Y \)

\[
J(y) := \inf \{ I(x) : x \in X, y = f(x) \}.
\]

is a rate function on \( Y \), where the infimum over the empty set is taken as \( \infty \).

(b) If \( \{X_n\} \) satisfies the Laplace principle on \( X \) with rate function \( I \), then \( \{f(X_n)\} \) satisfies the Laplace principle on \( Y \) with rate function \( J \).

Another important property of the Laplace principle is that it is preserved under superexponential approximation.

**Theorem 2.24** ([13]) Let \( X^n \) and \( Y^n \) be random variables defined on the same probability space \( (\Omega, \mathcal{F}, \mathbb{P}) \) and taking values in \( X \). Assume that \( \{X^n\} \) satisfies the Laplace principle on \( X \) with the rate function \( I \) and that \( \{Y^n\} \) is superexponentially closed (exponentially equivalent) to \( \{X^n\} \) in the following sense: for each \( \delta > 0 \)

\[
\limsup_{n \to \infty} \frac{1}{n} \log \mathbb{P}\{d(Y^n, X^n) > \delta\} = -\infty.
\]

Then \( \{Y^n\} \) satisfies the Laplace principle on \( X \) with the same rate function \( I \).

### 2.3.0.1 Examples

The role of large deviations in evaluating the difference between the time-averaged value and its expectations when this difference is significant is best understood through the following example. Suppose that \( X_1, X_2, ..., X_n, ... \) is a sequence...
of i.i.d. random variables, for instance, normally distributed with mean zero and variance 1. Then
\[ E[e^{\theta(X_1 + \cdots + X_n)}] = e^{n(\theta^2/2)}. \]

On the other hand, \[ E[e^{\theta(X_1 + \cdots + X_n)}] = E[e^{n\theta(S_n/n)}] \]
and since \( S_n/n \) converges to zero almost surely (the law of large numbers), it is further equal to \( E[e^{o(n)}] \) which is different from \( e^{o(n)} \).

Indeed, assuming that \( \theta > 0 \), for any \( a > 0 \),
\[ E[e^{\theta S_n}] \geq e^{n\theta a} P\left\{ \frac{S_n}{n} \geq a \right\} = e^{n\theta a} e^{-na^2/2+o(n)} = e^{n(a\theta-a^2/2)+o(n)}. \]

Since \( a > 0 \) is arbitrary chosen,
\[ E[e^{\theta S_n}] \geq e^{n\sup_{a>0}(a\theta-a^2/2)+o(n)} = e^{n\theta^2/2+o(n)}. \]

The simplest example for which one can calculate probabilities of large deviations is coin tossing ([14]).

1. Large deviations for sum of i.i.d. random variables

Consider a sequence of independent tosses of a fair coin. The probability of \( k \) heads in \( n \) tosses is \( P(n, k) = C_n^k 2^{-n} \approx \frac{\sqrt{2\pi} e^{-(n-k)(n-k)} (n-k)^{n-k}}{\sqrt{2\pi} e^{-k^2} n^{k+1/2} 2^{-n}} \).

Stirling’s approximation. Therefore, \( \log P(n, k) \approx -\frac{1}{2} \log(2\pi) - \frac{1}{2} \log n - (n-k+\frac{1}{2}) \log(1-\frac{k}{n}) - (k+\frac{1}{2}) \log \frac{k}{n} - n \log 2. \)

If \( k \approx nx \), then
\[ \log P(n, k) \approx -n[\log 2 + x \log x + (1-x) \log(1-x)] + o(n) \]
\[ = -nH(x) + o(n) \]
where $H(x)$ is the Kullback-Leibler information or relative entropy of $\text{Binomial}(x, 1 - x)$ with respect to $\text{Binomial}(\frac{1}{2}, \frac{1}{2})$.

Also, if $f_i$ are the observed frequencies in $n$ trials of a multinomial with probabilities $\{p_i\}$ for the individual cells, then

$$P(n, p_1, \ldots, p_k; f_1, \ldots, f_k) = \frac{n!}{f_1! \cdots f_k!} p_1^{f_1} \cdots p_k^{f_k}.$$

A similar calculation using Stirling’s approximation yields, assuming $f_i \approx nx_i$,

$$\log P(n, p_1, \cdots, p_k; f_1, \cdots, f_k) = -nH(x_1, \ldots, x_k; p_1, \ldots, p_k) + o(n)$$

where $H(x, p)$ is again the Kullback-Leibler information number

$$H(x) = \sum_{i=1}^{k} x_i \log \frac{x_i}{p_i}.$$

Any probability distribution can be approximated by one that is concentrated on a finite set and the empirical distribution from a sample of size $n$ will have a multinomial distribution. One therefore expects that the probability $P(n, \mu, \nu)$ that the empirical distribution $\frac{1}{n} \sum_{i=1}^{n} \delta_{X_i}$ of $n$ independent observations from a distribution $\mu$ is close to $\nu$ should satisfy

$$\log P(n, \mu, \nu) = -nH(\nu|\mu) + o(n)$$

where $H(\nu|\mu)$ is again the Kullback-Leibler information number

$$\int \log \frac{d\nu}{d\mu} d\nu = \int \frac{d\nu}{d\mu} \log \frac{d\nu}{d\mu} d\mu.$$ (2.3.2)
The quantity $H(\nu|\mu)$ is also called the relative entropy of $\nu$ with respect to $\mu$. This is rigorously proven by Sanov.

**Theorem 2.25 (Sanov [15])** Let $\mu$ be a probability measure on the Polish space $\mathcal{X}$. Let for $X = (X_1, ..., X_n) \in \mathcal{X}^n$ define $L_n(X) := \frac{1}{n} \sum_{m=1}^{n} \delta_{X_m}$. Let $\tilde{\mu}_n \in \mathcal{M}_1(\mathcal{M}_1(\mathcal{X}))$ be the distribution under $\mu^n$ of the function $L_n$. Then the family $\{\tilde{\mu}_n : n \geq 1\}$ satisfies the large deviation principle with rate function $I(\cdot) = H(\cdot|\mu)$ where

$$
H(\nu|\mu) = \begin{cases} 
\int f \log f \, d\mu & \text{if } \nu \ll \mu \text{ and } f = \frac{d\nu}{d\mu} \\
\infty & \text{otherwise}
\end{cases}
$$

is the relative entropy of $\nu$ with respect to $\mu$.

A classical example for which the rate function is evaluated is Cramér theorem.

Let $X_1, X_2, \ldots$ be i.i.d. $d$-dimensional random vectors with $X_1$ distributed according to the probability law $\mu \in \mathcal{M}_1(\mathbb{R}^d)$. Consider the empirical means $\hat{S}_n := \frac{1}{n} \sum_{i=1}^{n} X_i$. The *cumulant generating function* associated with the law $\mu$ is defined as

$$
\Lambda(\lambda) := \log \mathbb{E}[e^{<\lambda, X_1>}] \quad (2.3.4)
$$

where $<\lambda, x> := \sum_{j=1}^{d} \lambda_j x_j$ is the usual scalar product in $\mathbb{R}^d$.

Let $\mu_n$ be the law of $\hat{S}_n$ and $\bar{x} := \mathbb{E}[X_1]$. When $\bar{x}$ exists and is finite and $\mathbb{E}[|X_1 - \bar{x}|^2] < \infty$, then $\hat{S}_n$ converges in probability to $\bar{x}$. Therefore $\mu_n(F) \to 0$ as $n \to \infty$ for any closed set $F$ such that $\bar{x} \notin F$.

**Theorem 2.26 (Cramér’s theorem [12])** Assume that $\Lambda(\lambda) < \infty$ for all $|\lambda|$ small enough. Then the sequence of measures $\{\mu_n\}$ satisfies the LDP on $\mathbb{R}^d$, that is

$$
\lim_{n \to \infty} \frac{1}{n} \log \mu_n(A) = - \inf_{x \in A} \Lambda^*(x).
$$
where the rate function

$$\Lambda^*(x) := \sup_{\lambda \in \mathbb{R}^d} \{ \langle \lambda, x \rangle - \Lambda(\lambda) \}$$

is the Fenchel-Legendre transform of \( \Lambda(\lambda) \).

2. Sample path large deviations for random walk and Brownian motion

Let \( X_1, X_2, ... \) be a sequence of i.i.d. random variables taking values in \( \mathbb{R}^d \) with the cumulant generating function \( \Lambda(\lambda) \) defined in (2.3.4) being finite for all \( \lambda \in \mathbb{R}^d \). Let

$$Z_n(t) = \frac{1}{n} \left[ \sum_{i=1}^{[nt]} X_i \right], \quad 0 \leq t \leq 1,$$

and let \( \mu_n \) be the law of \( Z_n(\cdot) \). Then

**Theorem 2.27** (Mogulskii’s theorem [16]) The measures \( \mu_n \) satisfy the large deviation principle with rate function

$$I(\phi) = \begin{cases} \int_0^1 \Lambda^*(\phi'(t)) \, dt, & \text{if } \phi \in A_0[0, 1] \\ \infty & \text{otherwise} \end{cases}$$

where \( \Lambda^* \) is the Fenchel-Legendre transform of \( \Lambda(\cdot) \).

From this, another important result is derived, the large deviation principle for Brownian motion. Consider a standard Brownian motion \((B(t), t \in \mathbb{R}_+)\), and \( X_n(t) = \frac{B(t)}{\sqrt{n}} \), where \( n \) is a large parameter. Since \( X_n(t) - X_n(s) \) are Gaussian with mean zero and variance \( \frac{t-s}{n} \), the sequence of random variables verify LDP with rate function \( \frac{x^2}{2(t-s)} \). Let us now consider a vector \((X_n(t_i), i = 1, 2, ... k)\), where \( t_1 < t_2 ... < t_k \). By the property of independence of increments, the \( k \)-dimensional vector \((X_n(t_i) - X_n(t_{i-1}), i = 1, 2, ..., k)\), \( t_0 = 0 \), verify the LDP with rate function \( \frac{1}{2} \sum_{i=1}^{k} \frac{x_i^2}{t_i-t_{i-1}} \). Therefore, by the contraction principle, the
vectors \( (X_n(t_i), i = 1, 2, \ldots, k) \) verify the LDP with rate \( \frac{1}{2} \sum_{i=1}^{k} \frac{(x_i - x_{i-1})^2}{t_i - t_{i-1}} \). This means that for small enough \( \varepsilon > 0 \) and large \( n \)

\[
\frac{1}{n} \log \mathbb{P}\{|X_n(t_i) - x_i| \leq \varepsilon, i = 1, 2, \ldots, k\} \approx -\frac{1}{2} \sum_{i=1}^{k} \left( \frac{x_i - x_{i-1}}{t_i - t_{i-1}} \right)^2 (t_i - t_{i-1}).
\]

Therefore, if \((x(t), t \in [0, 1])\) is an absolutely continuous function, then by considering finer and finer subdivisions of the \([0, 1]\) one should expect that

\[
\frac{1}{n} \log \mathbb{P}\{|X_n(t) - x(t)| \leq \varepsilon\} \approx -\frac{1}{2} \int_0^1 x'(t)^2 \, dt,
\]

so that the rate function should be \( \frac{1}{2} \int_0^1 x'(t)^2 \, dt \) for absolutely continuous functions starting at zero. The right-hand side tends to \( -\infty \) if \( x \) is either not absolutely continuous or if it does not start at zero, rendering the action functional equal infinity in these cases.

**Theorem 2.28** (Schilder’s theorem [12]) Let \( \mathcal{A}_0([0, 1]) \) denote of all absolutely continuous functions \( \phi \) satisfying \( \phi(0) = 0 \). Then \( X_n(t) := \frac{1}{\sqrt{n}} B(t) \) satisfies the large deviation principle on \( \mathcal{C}[0, 1] \) with rate function

\[
I(\phi) = \begin{cases} 
\frac{1}{2} \int_0^1 \phi'(t)^2 \, dt, & \text{if } x \in \mathcal{A}_0([0, 1]) \\
\infty & \text{if } x \in \mathcal{C}([0, 1]) \setminus \mathcal{A}_0([0, 1])
\end{cases}
\]

This statement can be generalized to Itô diffusions by applying various contraction principles.

Let \( B \) be the standard Brownian motion and \( X_t^\varepsilon \) be the diffusion process, that is the unique solution of the stochastic differential equation

\[
dX_t^\varepsilon = b(X_t^\varepsilon) \, dt + \sqrt{\varepsilon} dB_t \quad 0 \leq t \leq 1, \quad X_0^\varepsilon = 0,
\]
where $b : \mathbb{R} \to \mathbb{R}$ is a uniformly continuous function.

**Theorem 2.29** (*Freidlin-Wentzell theorem [12]*) The family $\{X_t^\varepsilon\}$ satisfies the LDP in $\mathcal{C}_0([0, 1])$ with rate function

$$I(\phi) = \begin{cases} 
\frac{1}{2} \int_0^1 |\phi'(t) - b(\phi(t))|^2 \, dt, & \text{if } f \in \mathcal{A}_0([0, 1]) \\
\infty & \text{otherwise}
\end{cases}$$
The main mathematical object of this chapter is a family of coupled Markov process $(\xi(t), x(t)), t \geq 0$ called the switched and switching processes, respectively. The switched process describes the evolution of the system and it is a stochastic functional of the process $\eta(t; x), t \geq 0, x \in E$ with locally independent increments [5]. In order to reduce the complexity of the phase space, the switching processes that describe the random changes in the evolution of the system, are jump Markov processes considered in a split space $E = \bigcup_{k=1}^{N} E_k, E_k \cap E_{k'} = \emptyset, k \neq k'$ with non-communicating components, and having the ergodic property on each class $E_k$. By introducing the parameter $\epsilon > 0$ one defines a jump Markov process on the split phase space with small transition probabilities between the states of the system and further merges the classes $E_k, k = 1, 2, \cdots, N$ into distinct states $k, 1 \leq k \leq N$.

The average limit theorem of the stochastic additive functional with fast time-scaling switching process is obtained by using the martingale characterization [8] and a solution of the singular perturbation problem for reducible-invertible operators [4]. We are interested in finding the large deviation principle for this sequence of stochastic additive functionals. Using the weak convergence approach of Dupuis and Ellis [13], a large deviation principle is derived for a sequence of random walks constructed such that they have the same distribution as the linear interpolation sequence of samples of stochastic additive functionals.
3.1 Average approximation

3.1.1 Stochastic additive functionals

**Definition 3.1** An \( \mathbb{R}^d \times E \)-valued coupled stochastic process \((\xi(t), x(t)), t \geq 0\) is called a Markov additive process (see [17], [18] for details) if

(a) the coupled process \((\xi(t), x(t)), t \geq 0\) is a Markov process;

(b) on \(\{\xi(t) = u\}\) we have a.s. the following relation

\[ P(\xi(t+s) \in A, x(t+s) \in B | \mathcal{F}_t) = P(\xi(t+s) - \xi(t) \in A - u, x(t+s) \in B | x(t)) \]

**Example 3.2** Sums of i.i.d. random vectors in \( \mathbb{R}^d \).

**Example 3.3** A Markov renewal process \( \xi_n = \tau_n, x_n, n \geq 0 \) is a Markov additive process with additive component \( \xi_n \).

**Definition 3.4** The Markov additive process \((\xi(t), x(t)), t \geq 0\) in \( \mathbb{R}^d \times E \) with Markov switching \( x(t), t \geq 0 \), defined by the relation

\[ \xi(t) = \xi_0 + \int_0^t \eta(ds; x(s)), t \geq 0, \]

is called a stochastic additive functional.

The family of cadlag Markov processes \( \eta(t; x), t \geq 0, x \in E \), parameterized by \( x \), are such that \( \eta(t; x(t)) \) is measurable, are of locally independent increment processes determined by their infinitesimal generators

\[ \Gamma(x)\varphi(u) = a(u; x)\varphi'(u) + \int_{\mathbb{R}^d} \left[ \varphi(u + v) - \varphi(u) - v\varphi'(u) \right] \Gamma(u, dv; x) \]
where the positive kernels $\Gamma(u, dv; x), x \in E$, satisfy the following conditions: the functions

$$
\Lambda(u; x) := \Gamma(u, \mathbb{R}^d; x)
$$

$$
b(u; x) := \int_{\mathbb{R}^d} v \Gamma(u, dv; x)
$$

$$
C(u; v) := \int_{\mathbb{R}^d} vv^* \Gamma(u, dv; x)
$$

are continuous and bounded on $u \in \mathbb{R}^d$, and uniformly continuous and bounded on $x \in E$. The product $a \varphi'$ stands for the inner product $\langle a, \nabla \varphi \rangle$ in $\mathbb{R}^d$.

The switching jump Markov process $x(t), t \geq 0$ is defined by its infinitesimal generator $Q$ as

$$
Q \varphi(x) = q(x) \int_{E} P(x, dy)[\varphi(y) - \varphi(x)],
$$

where the kernel $P(x, B)$ is the transition kernel of the embedded Markov chain, and $q(x), x \in E$ is the intensity of jumps function.

**Lemma 3.5** The Markov additive process $(\xi(t), x(t)), t \geq 0$ is determined by the infinitesimal generator

$$
\mathbb{L} \varphi(u, x) = Q \varphi(u, x) + \Gamma(x) \varphi(u, x).
$$
Proof: We have:

\[
\mathbb{E}[\varphi(u(t + \Delta t), x(t + \Delta t))|\xi(t) = u, x(t) = x] = \\
\mathbb{E}[\varphi(u + \Delta \xi(t), x(t + \Delta t))|\xi(t) = u, x(t) = x] = \\
\mathbb{E}[\varphi(u + \Delta \xi(t), x(t + \Delta t))|x(t) = x] = \\
\mathbb{E}[\varphi(u + \Delta \xi(t), x(t + \Delta t))\mathbb{1}_{\{\theta_{x} > \Delta t\}}|x(t) = x] + \\
\mathbb{E}[\varphi(u + \Delta \xi(t), x(t + \Delta t))\mathbb{1}_{\{\theta_{x} < \Delta t\}}|x(t) = x] + o(\Delta t)
\]

where \(\theta_{x}\) is the sojourn time at \(x\). It follows that the above it is further equal to

\[
\mathbb{E}[\varphi(u + \Delta \eta(t; x), x)](1 - \Delta tQ(x)) + \mathbb{E}[\varphi(u, x(t + \Delta t))|\Delta tQ(x) + \\
\varphi'(u)\Delta \eta(t; x)\Delta tQ(x) + o(\Delta t) = \mathbb{E}[\varphi(u + \Delta \eta(t; x), x)] + \Delta tQ\varphi(u; x) + o(\Delta t) \\
= \varphi(u, x) + \Gamma(x)\varphi(u, x)\Delta t + \Delta tQ\varphi(u, x) + o(\Delta t) \\
= \varphi(u, x) + \Delta t[\Gamma(x) + Q]\varphi(u, x) + o(\Delta t)
\]

which ends the proof.

3.1.2 Phase merging principle

The general scheme of phase merging is realized by the family of time-homogeneous cadlag Markov jump process \(x'(t), t \geq 0, \epsilon > 0\) with the standard phase space \((E, \mathcal{E})\), on the split phase space

\[E = \bigcup_{k=1}^{N} E_k, \quad E_k \bigcap E_{k'} = \emptyset, \quad k \neq k'\]
given by the infinitesimal generator

$$Q^\epsilon \varphi(x) = q(x) \int_E P^\epsilon(x,dy) [\varphi(y) - \varphi(x)].$$

(3.1.5)

The phase merging algorithm is considered under the following assumptions:

A1. The stochastic kernel in (3.1.5) is represented in the following form

$$P^\epsilon(x,B) = P(x,B) + \epsilon P_1(x,B)$$

where the stochastic kernel $P(x,B)$ is coordinated with the splitting as follows:

$$P(x,E_k) = \mathbb{I}_k(x) := \begin{cases} 1, & x \in E_k, \\ 0, & x \notin E_k \end{cases}$$

A2. The Markov supporting process $x(t), t \geq 0$ on the state space $(E, \mathcal{E})$, determined by the generator $Q$ given in (3.1.3) is supposed to be uniformly ergodic in every class $E_k, 1 \leq k \leq N$, with the stationary distribution $\pi_k(dx), 1 \leq k \leq N$, satisfying the following relations

$$\pi_k(dx)q(x) = q_k \rho_k(dx), \quad q_k = \int_{E_k} \pi_k(dx)q(x),$$

$$\rho_k(B) = \int_{E_k} \rho_k(dx)P(x,B), \quad \rho_k(E_k) = 1.$$ 

The perturbing operator $P_1(x,B)$ is a signed kernel which satisfies the conservative condition $P_1(x,E) = 0$. 
A3. The average exit probabilities satisfy the following condition

\[ \hat{p}_k := \int_{E_k} \rho_k(dx)P_1(x,E \setminus E_k) > 0, \quad 1 \leq k \leq N. \]

Introduce the merging function \( m(x) = k, x \in E_k, 1 \leq k \leq N, \) and the merged process

\[ \hat{x}^\epsilon(t) := m(x^\epsilon(t/\epsilon)), \quad t \geq 0 \quad (3.1.6) \]

on the merged phase space \( \hat{E} = \{1, \ldots, N\}. \)

The phase merging principle establishes the weak convergence of the above process to the limit Markov process \( \hat{x}(t) \) (a variant of Theorem 4.1 in [5]).

**Theorem 3.6 (Ergodic phase merging principle)** Under the assumptions \( A1 - A3, \) the following weak convergence holds

\[ \hat{x}^\epsilon(t) \Rightarrow \hat{x}(t) \quad \text{as} \quad \epsilon \to 0. \]

The limit merged Markov process \( \hat{x}(t), t \geq 0, \) on the merged state space \( \hat{E} \) is determined by the generator matrix

\[ \hat{Q} = (\hat{q}_{kr}; 1 \leq k, r \leq N), \]

with entries

\[ \hat{q}_{kr} = \hat{q}_kp_{kr}, \quad p_{kr} = \int_{E_k} \rho_k(dx)P_1(x,E_r), \quad 1 \leq k, r \leq N \quad (3.1.7) \]

where \( \rho_k \) is the stationary distribution of the corresponding embedded Markov chain.
Proof: Since $x^\epsilon(t)$ is a jump Markov process with infinitesimal generator $Q+\epsilon Q_1$, then $x^\epsilon(\frac{t}{\epsilon})$ has the generator $Q^\epsilon := \epsilon^{-1}Q + Q_1$.

Using the martingale characterization of $x^\epsilon(\frac{t}{\epsilon})$ we have that

$$
\mu^\epsilon_t := \varphi^\epsilon \left( x^\epsilon \left( \frac{t}{\epsilon} \right) \right) - \int_0^t Q^\epsilon \varphi^\epsilon \left( x^\epsilon \left( \frac{s}{\epsilon} \right) \right) ds
$$

with $\varphi^\epsilon \in D(Q^\epsilon)$, is a martingale.

Let us consider the test functions of the form $\varphi^\epsilon = \varphi + \epsilon \varphi_1$. $Q$ is a reducible-invertible operator, so using Proposition 2.11 the asymptotic representation

$$(\epsilon^{-1}Q + Q_1)(\varphi + \epsilon \varphi_1) = \psi + \epsilon \theta_\epsilon$$

is realized by

$$
\hat{Q}_1 \hat{\varphi} = \hat{\psi} \\
\varphi_1 = R_0(\psi - Q_1 \varphi) \\
\theta_\epsilon = Q_1 R_0(\psi - Q_1 \varphi) \\
R_0 = (Q + \Pi)^{-1} - \Pi \\
\Pi Q_1 \Pi = \hat{Q}_1 \Pi
$$

Let us define $\hat{\varphi}: \hat{E} \rightarrow \mathbb{R}$ by $\hat{\varphi}(m(x)) = \varphi(x)$. Then $\varphi^\epsilon = \hat{\varphi}(m(x)) + \epsilon \varphi_1(x)$ and

$$(\epsilon^{-1}Q + Q_1)(\varphi + \epsilon \varphi_1) = \hat{Q}_1 \hat{\varphi}(m(x)) + \epsilon \theta_\epsilon.$$
This implies that

\[ \mu_t^\epsilon := \dot{\phi} \left( m(x^\epsilon(t/\epsilon)) \right) + \epsilon \varphi_1 \left( x^\epsilon(t/\epsilon) \right) - \int_0^t \hat{Q}_1 \dot{\phi} \left( m(x^\epsilon(s/\epsilon)) \right) \, ds - \int_0^t \epsilon \theta^\epsilon \left( x^\epsilon(s/\epsilon) \right) \, ds \]

is a martingale. Therefore

\[ \dot{\phi} \left( m(x^\epsilon(t/\epsilon)) \right) - \int_0^t \hat{Q}_1 \dot{\phi} \left( m(x^\epsilon(s/\epsilon)) \right) \, ds = \mu_t^\epsilon + \epsilon \psi_t^\epsilon, \]

where

\[ \psi_t^\epsilon = \int_0^t \theta^\epsilon \left( x^\epsilon(s/\epsilon) \right) \, ds - \varphi_1(x^\epsilon_t). \]

Applying Theorem 2.15 we get that \( m \left( x^\epsilon(t/\epsilon) \right) \Rightarrow \hat{x}(t) \) and \( \hat{x}(t) \) is the solution of the martingale problem

\[ \dot{\mu}_t = \dot{\phi}(\hat{x}(t)) - \int_0^t \hat{Q}_1 \dot{\phi}(\hat{x}(s)) \, ds \]

with infinitesimal generator \( \hat{Q}_1 \) determining the merged Markov process \( \hat{x}(t) \) on the merged phase space \( \hat{E} \).

The contracted operator \( \hat{Q}_1 \) is defined by \( \hat{Q}_1 \Pi = \Pi Q_1 \Pi \). Therefore,

\[ \Pi Q_1 \Pi \varphi(x) = \Pi Q_1 \sum_{k=1}^N \hat{\phi}_k \Pi_k(x) = \sum_{k=1}^N \hat{\phi}_k \Pi \int_{E_k} Q_1(x, dy) \]

\[ = \sum_{k=1}^N \hat{\phi}_k \Pi Q_1(x, E_k) = \sum_{k=1}^N \hat{\phi}_k \sum_{r=1}^N \hat{q}_{rk} \Pi_r(x) \]

where \( \hat{q}_{rk} := \int_{E_r} \pi_r(dx) Q_1(x, E_k) \), \( 1 \leq r, k \leq N \). Thus

\[ \Pi Q_1 \Pi \varphi(x) = \sum_{r=1}^N \Pi_r(x) \sum_{k=1}^N \hat{q}_{rk} \hat{\phi}_k \]
and the contracted operator $\hat{Q}_1$ is determined by the matrix

$$
\hat{Q}_1 := (\hat{q}_{rk}; 1 \leq r, k \leq N)
$$

$$
\hat{Q}_1 \hat{\phi} = \hat{\psi}, \quad \hat{\psi}_r := \sum_{k=1}^{N} \hat{q}_{rk} \hat{\phi}_k, \quad 1 \leq r \leq N.
$$

Let us consider the stochastic additive functionals $\xi^\varepsilon(t), t \geq 0$ in the following scheme:

$$
\xi^\varepsilon(t) = \xi^\varepsilon(0) + \int_0^t \eta^\varepsilon \left( ds; x^\varepsilon(t) \right), \quad t \geq 0, \varepsilon > 0. \quad (3.1.8)
$$

where the process $\eta^\varepsilon(t; x), t \geq 0, \varepsilon > 0, x \in E$ is given by the infinitesimal generators

$$
\Gamma^\varepsilon(x) \varphi(u) = a^\varepsilon(u; x) \varphi'(u) + \varepsilon^{-1} \int_{\mathbb{R}^d} [\varphi(u + \varepsilon v) - \varphi(u) - \varepsilon v \varphi'(u)] \Gamma_c(u, dv; x). \quad (3.1.9)
$$

The Markov additive process $(\xi^\varepsilon(t), x^\varepsilon(t)), t \geq 0, \varepsilon > 0$ is determined by the infinitesimal generator

$$
\mathbb{L}^\varepsilon \varphi(u, x) = Q^\varepsilon \varphi(u, x) + \Gamma^\varepsilon(x) \varphi(u, x). \quad (3.1.10)
$$

We want to prove the weak convergence of the stochastic system with Markov switching in the series scheme, $(\xi^\varepsilon(t), x^\varepsilon(t)), t \geq 0, \varepsilon > 0$. For this we use the singular perturbation problem for reducible-invertible operators and the martingale characterization approach of Stroock and Varadhan to find the relative compactness of the process. The coupled Markov process is characterized by the martingale

$$
\mu^\varepsilon(t) = \varphi(\xi^\varepsilon(t), x^\varepsilon(t)) - \int_0^t \mathbb{L}^\varepsilon \varphi(\xi^\varepsilon(s), x^\varepsilon(s)) ds
$$
where the generators $L^\varepsilon$, $\varepsilon > 0$ have the common domain of definition $\mathcal{D}(L)$ supposed to be dense in $C^2_0(\mathbb{R}^d \times E)$.

**Lemma 3.7** Let the generators $L^\varepsilon$, $\varepsilon > 0$ have the following property

$$|L^\varepsilon \varphi(u)| \leq c_\varphi$$  \hspace{1cm} (3.1.11)

for any real-valued nonnegative function $\varphi \in C^2_0(\mathbb{R}^d)$, where the constant $c_\varphi$ depend only on the norm of $\varphi$.

Suppose that the compact containment condition holds

$$\lim_{l \to \infty} \sup_{\varepsilon > 0} \mathbb{P}( \sup_{0 \leq t \leq T} |\xi^\varepsilon(t)| > l) = 0$$  \hspace{1cm} (3.1.12)

Then the family of stochastic processes $\xi^\varepsilon(t)$, $t \geq 0$, $\varepsilon > 0$ is relatively compact.

**Proof:** Let us consider the process $\eta^\varepsilon(t) := \varphi(\xi^\varepsilon(t)) + c_\varphi t$, $t \geq 0$, where $\varphi$ is understood to be composed with the projection on the first component. We want to prove that $\eta^\varepsilon(t)$ is an $\mathcal{F}_t^\varepsilon$-submartingale. For $0 \leq s \leq t$ we have

$$\mathbb{E}[\eta^\varepsilon(t)|\mathcal{F}_s^\varepsilon] = \mathbb{E} \left[ \int_s^t L^\varepsilon \varphi(\xi^\varepsilon(u)) \, du |\mathcal{F}_s^\varepsilon \right] + \mu_s^\varepsilon + c_\varphi s + \mathbb{E} \left[ \int_s^t (L^\varepsilon \varphi(\xi^\varepsilon(u)) + c_\varphi) \, du |\mathcal{F}_s^\varepsilon \right]$$

$$= \eta^\varepsilon(s) + \mathbb{E} \left[ \int_s^t (L^\varepsilon \varphi(\xi^\varepsilon(u)) + c_\varphi) \, du |\mathcal{F}_s^\varepsilon \right]$$

where the last term is nonnegative due to the condition (3.1.11). So the conditions of Theorem 2.14 are fulfilled, and therefore $\xi^\varepsilon(t)$, $t \geq 0$, $\varepsilon > 0$ is relatively compact.
Lemma 3.8 Let the generators $\mathbb{L}^\varepsilon$, $\varepsilon > 0$ have the following estimation for $\varphi_0(u) = \sqrt{1 + u^2}$,

$$\mathbb{L}^\varepsilon \varphi_0(u) \leq c_0 \varphi_0(u), \quad |u| \leq l_0$$

where the constant $c_0$ depends on the function $\varphi_0$, but not on $\varepsilon > 0$, and $\mathbb{E}[|\xi^\varepsilon(0)|] \leq b < \infty$. Then the compact containment condition (3.1.12) holds.

Proof: $\varphi_0'(u) = \frac{u}{\sqrt{1 + u^2}}$, $\varphi_0''(u) = \frac{1}{(1 + u^2)\sqrt{1 + u^2}}$. Then $\varphi_0(u) \leq 1 + |u|$, $|\varphi_0'(u)| \leq 1 \leq \varphi_0(u)$, $|\varphi_0''(u)| \leq 1 \leq \varphi_0(u)$.

It is known (Lemma 3.2 [9]) that if $x(t), t \geq 0$ is a Markov process defined by the generator $\mathbb{L}$ and if $\varphi \in \mathcal{D}(\mathbb{L})$ then

$$e^{-\lambda t} \varphi(x(t)) - \int_0^t e^{-\lambda s}[\lambda \varphi(x(s)) - \mathbb{L} \varphi(x(s))] ds$$

is a martingale.

Let us define the stopping time $\tau^\varepsilon_t$ by

$$\tau^\varepsilon_t := \begin{cases} \inf \{t \in [0, T] : |\xi^\varepsilon(t)| \geq l \} \\ T \quad \text{if } |\xi^\varepsilon(t)| \leq l \text{ for all } t \in [0, T] \end{cases}$$

Then

$$e^{-c_0 t \wedge \tau^\varepsilon_t} \varphi_0(\xi^\varepsilon(t \wedge \tau^\varepsilon_t)) + \int_0^{t \wedge \tau^\varepsilon_t} e^{-c_0 s}[c_0 \varphi_0(\xi^\varepsilon(s)) - \mathbb{L}^\varepsilon \varphi_0(\xi^\varepsilon(s))] ds$$

is a martingale. We get, for $s \leq t \wedge \tau^\varepsilon_t$, $c_0 \varphi_0(\xi^\varepsilon(s)) - \mathbb{L}^\varepsilon \varphi_0(\xi^\varepsilon(s)) \geq 0$, so

$$\mathbb{E} \left[ e^{-c_0 t \wedge \tau^\varepsilon_t} \varphi_0(\xi^\varepsilon(t \wedge \tau^\varepsilon_t)) \right] \leq \mathbb{E} \mu^\varepsilon_t = \mathbb{E} \mu^\varepsilon_0 = \mathbb{E} \varphi_0(\xi^\varepsilon_0).$$
Then

$$\mathbb{P}^\varepsilon_t := \mathbb{P}^\varepsilon(\sup_{0 \leq t \leq T} |\xi^\varepsilon(t)| > l) \leq \mathbb{P}^\varepsilon(\varphi_0(\xi^\varepsilon(\tau^\varepsilon_T)) \geq \varphi_0(l))$$

$$\leq \frac{\mathbb{E}[\varphi_0(\xi^\varepsilon(\tau^\varepsilon_T))]}{\varphi_0(l)} \leq e^{c_0 T} \frac{\mathbb{E}[\varphi_0(\xi^\varepsilon(0))]}{\varphi_0(l)}$$

$$\leq e^{c_0 T} \frac{1 + \mathbb{E}[\xi^\varepsilon(0)]}{\varphi_0(l)} \leq e^{c_0 T} \frac{(1 + b)}{\varphi_0(l)}$$

which converges to 0 as $l \to \infty$.

**Corollary 3.9** Let the generators $\mathbb{I}^\varepsilon$, $\varepsilon > 0$ have the following estimation:

$$|\mathbb{I}^\varepsilon \varphi(u)| \leq c_\varphi \text{ for any real-valued nonnegative function } \varphi \in C^2_0(\mathbb{R}^d), \text{ where the constant } \varphi \text{ depends only on the norm of } \varphi, \text{ and for } \varphi_0(u) = \sqrt{1 + u^2}, \mathbb{I}^\varepsilon \varphi_0(u) \leq c_{\varphi_0} \varphi_0(u),$$

$$|u| \leq l_0, \text{ with the constant } c_{\varphi_0} \text{ depending only on the function } \varphi_0, \text{ but not on } \varepsilon > 0.$$

Then, the family of processes $\xi^\varepsilon(t), t \geq 0, \varepsilon > 0$ is relatively compact.

Let the switching Markov process $x^\varepsilon(t), t \geq 0$ satisfies the phase merging condition of Theorem 3.6. Let the following conditions be valid

C1. the drift velocity $a(u; x)$ belongs to the Banach space $B^1(\mathbb{R}^d)$, with

$$a^\varepsilon(u; x) = a(u; x) + \theta^\varepsilon(u; x)$$

where $\theta^\varepsilon(u; x) \to 0$ as $\varepsilon \to 0$ uniformly on $(u; x)$ and $\Gamma^\varepsilon(u, dv; x) \equiv \Gamma(u, dv; x)$ independent of $\varepsilon$.

C2. the operator $\gamma^\varepsilon(x) \varphi(u) = \varepsilon^{-1} \int_{\mathbb{R}^d} [\varphi(u + \varepsilon v) - \varphi(u) - \varepsilon v \varphi'(u)] \Gamma(u, dv; x)$ is negligible on $B^1(\mathbb{R}^d)$:

$$\sup_{\varphi \in C^1(\mathbb{R}^d)} ||\gamma^\varepsilon(x) \varphi|| \to 0 \quad \text{as} \quad \varepsilon \to 0$$
C3. the convergence in probability of the initial values of \((\xi^\epsilon(t), m(x^\epsilon(\xi)))\), \(t \geq 0\) holds, that is

\[(\xi^\epsilon(0), m(x^\epsilon(0)) \to (\xi(0), \hat{x}(0))\]

and there exists a constant \(c \in \mathbb{R}_+\) such that \(\sup_{\epsilon>0} \mathbb{E}|\xi^\epsilon(0)| \leq c < \infty\).

As in [19], under Markov switching, we get the following averaging theorem.

**Theorem 3.10 (Average approximation)** The stochastic evolutionary system \(\xi^\epsilon(t), t \geq 0\) defined by (3.1.8) converges weakly to the averaged stochastic system \(\hat{\xi}(t)\),

\[\xi^\epsilon(t) \Rightarrow \hat{\xi}(t) \quad \text{as} \quad \epsilon \to 0.\]

The limit process \(\hat{\xi}(t), t \geq 0\) is defined by a solution of the evolutionary equation

\[
\frac{d}{dt} \hat{\xi}(t) = \hat{a}(\hat{\xi}(t); \hat{x}(t)), \quad \hat{\xi}(0) = \xi(0),
\]

(3.1.13)

where the averaged velocity is determined by

\[\hat{a}(u; k) = \int_{E_k} \pi_k(dx)a(u; x) \quad 1 \leq k \leq N.\]

**Note 3.11** The limit process \(\hat{\xi}(t)\) is a random dynamical system evolving deterministically on random time intervals \([\tau_i, \tau_{i+1})\), where \(\{\tau_i\}_{i=1}^{N(T)}\) are the transition times of the stationary merged process \(\hat{x}(t)\) and \(N(T)\) the number of transitions on \([0, T]\).
Proof: First we will prove that the family \((\xi^\varepsilon(t), t \geq 0, \varepsilon > 0)\) is tight. For this, let’s consider the test functions \(\varphi_0 : \mathbb{R} \to [1, \infty), \varphi_0(u) = \sqrt{1 + u^2}\). We have 

\[
|\varphi_0'(u)| \leq 1 \leq \varphi_0(u), |\varphi_0''(u)| \leq 1 \leq \varphi_0(u), u \in \mathbb{R}.
\]

Then

\[
\mathbb{L}^\varepsilon \varphi_0(u) = \varepsilon^{-1}Q \varphi_0(u) + Q_1 \varphi_0(u) + \Pi(x) \varphi_0(u)
\]

\[
= \alpha^\varepsilon(u; x) \varphi_0'(u) \leq |\alpha^\varepsilon(u; x)| |\varphi_0(u)| \leq c_\alpha \varphi_0(u).
\]

By Lemma (3.8) the compact containment condition follows.

For \(\varphi \in C_0(\mathbb{R})\) we write:

\[
|\mathbb{L}^\varepsilon \varphi(u)| \leq q \varphi(u) + |\alpha^\varepsilon(u; x)| |\varphi'(u)| \leq c_{\alpha} \varphi(u)
\]

so using Lemma (3.7) it follows that \((\xi^\varepsilon(t), t \geq 0, \varepsilon > 0)\) is relatively compact and therefore tight in \(D([0, T])\) for any \(T > 0\).

The generator of the coupled Markov process \((\xi^\varepsilon(t), x^\varepsilon(\xi^\varepsilon(t)))\), \(t \geq 0\) is

\[
\mathbb{L}^\varepsilon = \varepsilon^{-1}Q + Q_1 + \Pi(x) + \gamma^\varepsilon(x) + \Sigma^\varepsilon(x)
\]

where \(\Pi(x) \varphi(u) = a(u; x) \varphi'(u)\), \(\Sigma^\varepsilon(x) \varphi(u) = \theta^\varepsilon(u; x) \varphi'(u)\) are bounded operators and the operator \(\gamma^\varepsilon(x) + \Sigma^\varepsilon(x)\) is negligible on \(B^1(\mathbb{R}^d)\).

Let \(R_0\) be the potential of the operator \(Q\), \(R_0 = [Q + \Pi]^{-1} - \Pi\).

Let \(\mathcal{C}_0^2(\mathbb{R}^d \times \hat{E})\) be the space of measurable, bounded functions \(\varphi(u, v)\) with compact support and twice continuously differentiable on the first argument.

We need to find the uniform convergence of generators, i.e.,

\[
\lim_{\varepsilon \to 0} \mathbb{L}^\varepsilon \varphi^\varepsilon(u, x) = \mathbb{L} \varphi(u, m(x)).
\]
Using the reducible-invertible operator method we have the following asymptotic representation

\[ [\varepsilon^{-1}Q + Q_1 + \Gamma(x)]\varphi^\varepsilon(u, x) = \mathbb{I}_L\hat{\varphi} + \varepsilon\Theta^\varepsilon(x) \]

with \( \varphi^\varepsilon(u, x) = \varphi(u, m(x)) + \varepsilon\varphi_1(u, x) \) and \( \hat{\varphi} = \hat{\varphi}(u, v) \in \tilde{\mathcal{C}}_0^2(\mathbb{R}^d \times \hat{E}) \), is realized by

\[ \hat{L} = \hat{Q}_1 + \hat{\Gamma}(x) \]

where contracted operators \( \hat{Q}_1 \) and \( \hat{\Gamma}(x) \) are defined by

\[
\begin{align*}
\Pi Q_1 \Pi &= \hat{Q}_1 \Pi \\
\Pi \Gamma(x) \Pi &= \hat{\Gamma}(x) \Pi \\
\varphi_1 &= R_0(\mathbb{I}_L - Q_1)\varphi \\
\Theta^\varepsilon &= (Q_1 + \Gamma(x))\varphi_1(u, x)
\end{align*}
\]

hence \( \Theta^\varepsilon(x) \) is a bounded operator independent of \( \varepsilon \) so \( \varepsilon \Theta^\varepsilon \to 0 \).

Indeed,

\[
\begin{align*}
[\varepsilon^{-1}Q + Q_1 + \Gamma(x)][\varphi(u, m(x)) + \varepsilon\varphi_1(u, x)] &= \\
\varepsilon^{-1}Q\varphi(u, m(x)) + \varepsilon[Q_1\varphi_1(u, x) + \Gamma(x)\varphi_1(u, x)] + Q\varphi_1(u, x) + Q_1\varphi(u, m(x))
\end{align*}
\]

From the condition \( Q\varphi(u, m(x)) = 0 \) we get that \( \varphi \in N_Q \) so \( \varphi = \Pi\varphi \). For

\[
Q\varphi_1(u, x) = \mathbb{I}_L\hat{\varphi} - Q_1\varphi(u, m(x)) - \Gamma(x)\varphi(u, m(x))
\]
under the solvability condition

\[ \Pi(\mathbb{L}\varphi - Q_1\varphi - \Gamma(x)\varphi) = 0 \]

which means

\[ \mathbb{L}\hat{\varphi} = (\hat{Q}_1 + \hat{\Gamma}(x))\hat{\varphi}, \]

we get the solution

\[ \varphi_1(u, x) = R_0(\mathbb{L} - Q_1)\varphi(u, m(x)) \]

Therefore \( \lim_{\varepsilon \to 0} \mathbb{L}^\varepsilon \varphi^\varepsilon(u, x) = \mathbb{L}\varphi(u, m(x)) \) uniformly, where

\[ \mathbb{L}\varphi(u, x) = \hat{Q}_1\hat{\varphi}(\cdot, x) + \hat{\Gamma}(x)\hat{\varphi}(u, \cdot). \]

It remains to calculate \( \hat{\Gamma}(x) \). Since \( \Pi(\Gamma(x))\varphi(u, \cdot) = a(u; x)\varphi'(u, \cdot) \) we get

\[ \hat{\Gamma}(x)\hat{\varphi}(u, \cdot) = \Pi\Gamma(x)\Pi\hat{\varphi}(u, \cdot) = \Pi a(u; x)\hat{\varphi}'(u, \cdot) = \hat{a}(u; x)\hat{\varphi}'(u, \cdot) \]

where

\[ \hat{a}(u; k) = \int_{E_k} \pi_k(dx)a(u; x) \]

Thus the limit process \( \hat{\xi}(t), t \geq 0 \) is defined by a solution of the evolutionary equation

\[ \frac{d}{dt}\hat{\xi}(t) = \hat{a}(\hat{\xi}(t); \hat{x}(t)), \quad \hat{\xi}(0) = \xi(0) \]

which ends the proof.
3.2 Large deviation principle for ergodic Markov processes

Let \( x(t), t \in \mathbb{R}_+ \) be a time-homogeneous Markov process on a compact metric space \( X, \mathcal{B}_X \) be the Borel \( \sigma \)-algebra in \( X \) and \( \mathcal{M}(X) \) the space of probability measures on \( \mathcal{B}_X \). Let introduce a random measure on \( \mathcal{B}_X \) by

\[
\nu_t(B) = \frac{1}{t} \int_0^t 1_{\{x(s) \in B\}} ds, \quad B \in \mathcal{B}_X.
\]

**Theorem 3.12** ([20] and [21]) Assume that the process \( x(t), t \in \mathbb{R}_+ \) is an ergodic Markov process. Then the following large deviation result holds

\[
-\inf_{m \in \Gamma} I(m) \leq \liminf_{t \to \infty} \frac{1}{t} \log P\{\nu_t \in \Gamma\} \leq \limsup_{t \to \infty} \frac{1}{t} \log P\{\nu_t \in \Gamma\} \leq -\inf_{m \in \bar{\Gamma}} I(m)
\]

where the rate function \( I : \mathcal{M}(X) \to [0, +\infty] \) is defined by

\[
I(m) = -\inf \left\{ \int (\phi(x))^{-1} Q\phi(x)m(dx) : \phi \in \mathcal{D}(Q), \phi > 0 \right\}
\]

and \( \Gamma \in \mathcal{B}(\mathcal{M}(X)) \) be the Borel \( \sigma \)-algebra in \( \mathcal{M}(X) \).

Typically \( \Gamma = \{ \nu \in \mathcal{M}(X) | d(\nu, m) > \delta, I(m) = 0 \} \) where \( d \) is some metric on \( \mathcal{M}(X) \).

The rate function \( I(m) \) verifies the following properties

(i) \( I(m) \geq 0 \) for all \( m \in \mathcal{M}(X) \), and \( I(m) = 0 \) if and only if \( m \) is the invariant measure for the ergodic Markov process,

(ii) \( I(m) \) is a convex function, i.e.,

\[
I(sm_1 + (1-s)m_2) \leq sI(m_1) + (1-s)I(m_2), \quad m_i \in \mathcal{M}(X), i = 1, 2, 0 < s < 1
\]
(iii) $I(m)$ is a lower semi-continuous function, i.e.,

$$\lim \inf_{m_n \to m} I(m_n) \geq I(m)$$

(iv) For any $b > 0$ the set $C_b(I) = \{ m : I(m) \leq b \}$ is compact, and the function $I(m)$ is continuous on this compact set.

We illustrate the concept of the split phase space in the following example.

**Example 3.13**

Let us consider a four-state Markov process $x(t), t \in \mathbb{R}_+$ on the split phase space $E = \{1, 2, 3, 4\} = E_1 \cup E_2, E_1 = \{1, 2\}, E_2 = \{3, 4\}$ generated by

$$Q = \begin{pmatrix}
-\lambda_1 & \lambda_1 & 0 & 0 \\
\mu_1 & -\mu_1 & 0 & 0 \\
0 & 0 & -\lambda_2 & \lambda_2 \\
0 & 0 & \mu_2 & -\mu_2
\end{pmatrix}$$

One checks that the Markov process $x(t)$ is ergodic in both $E_1$ and $E_2$ with stationary distributions $\pi_1 = (\frac{\mu_1}{\lambda_1 + \mu_1}, \frac{\lambda_1}{\lambda_1 + \mu_1})$ and $\pi_2 = (\frac{\mu_2}{\lambda_2 + \mu_2}, \frac{\lambda_2}{\lambda_2 + \mu_2})$.

Now we analyze singularly perturbed Markov processes by introducing a small parameter $\varepsilon > 0$ which leads to a singular perturbed system involving two-time scales, the actual time $t$ and the stretched time $\frac{t}{\varepsilon}$. Since the process $x(t)$ is ergodic on $E_1, E_2$, the system can be decomposed and the states of the Markov process can be aggregated.
Let $x^\epsilon(t)$ be a Markov chain on $E$ generated by $Q + \epsilon Q_1$ with $Q$ defined above and $Q_1$ given by

$$Q_1 = \begin{pmatrix} -\lambda_1 & 0 & \lambda_1 & 0 \\ 0 & -\mu_1 & 0 & \mu_1 \\ \lambda_2 & 0 & -\lambda_2 & 0 \\ 0 & \mu_2 & 0 & -\mu_2 \end{pmatrix}$$

and $x^\epsilon(\frac{t}{\epsilon})$ be a time-invariant Markov process with generator $Q^\epsilon = \frac{1}{\epsilon}Q + Q_1$.

Note that for small $\epsilon$, the Markov process $x^\epsilon(\frac{t}{\epsilon})$ jumps more frequently within each block and less frequently from one block to another. To further understanding of the underlying process, we consider the merged process $\hat{x}^\epsilon(t) := m(x^\epsilon(\frac{t}{\epsilon}))$ obtained by aggregating the states in the $k^{th}$ block by a single state $k$ and study its asymptotic behavior (for many asymptotic results see [22]).

Theorem 3.6 states that the limit process is a Markov process on the merged space $\hat{E} = \{1, 2\}$ determined by generator matrix $\hat{Q} = (\hat{q}_{kr}, 1 \leq k, r \leq 2)$ with $\hat{q}_{kr}$ verifying (3.1.7). For this example, $q_1 = \frac{2\lambda_1\mu_1}{\lambda_1+\mu_1}$, $q_2 = \frac{2\lambda_2\mu_2}{\lambda_2+\mu_2}$, $p_{11} = -1$, $p_{12} = 1$, $p_{21} = 1$, $p_{22} = -1$.

Thus,

$$\hat{Q} = \begin{pmatrix} \frac{2\lambda_1\mu_1}{\lambda_1+\mu_1} & \frac{2\lambda_2\mu_2}{\lambda_2+\mu_2} \\ \frac{2\lambda_2\mu_2}{\lambda_2+\mu_2} & \frac{2\lambda_1\mu_1}{\lambda_1+\mu_1} \end{pmatrix} := \begin{pmatrix} -\lambda & \lambda \\ \mu & -\mu \end{pmatrix}$$

Note 3.14 The merged process $\hat{x}^\epsilon(t)$, unlike its limit $\hat{x}(t)$, is not time-homogeneous.

Let us consider now the occupational time of $\hat{x}(t)$ defined by

$$\nu_t(B) = \frac{1}{t} \int_0^t \mathbb{1}_{\{\hat{x}(s) \in B\}}(s)ds$$

for any $B \in \mathcal{B}_\hat{E}$. By ergodic theorem, the measure $\nu_t$ converges to the ergodic distribution $\rho$ as $t$ goes to $\infty$. As an example, let $\mathcal{M}$ be the set of all probability
measures on \(\{0, 1\}\) identified with \(\{(p, 1-p), 0 \leq p \leq 1\}\) and \(d(x, y) = |x| + |y|, x, y \in \mathbb{R}^2\). Then Theorem 3.12 implies that for \(\rho = (p_0, 1-p_0)\) and \(\Gamma = \{(p, 1-p) | d((p, 1-p), (p_0, 1-p_0)) > \delta\}\), \(\mathbb{P}(\nu_t \in \Gamma) \sim \exp(-tI(p_0 + \delta, 1-p_0 - \delta))\) for large \(t\) with

\[
I(m) = -\inf \{ \int_E \frac{(\hat{Q}\phi(y))m(dy)}{\phi(y)} : \phi \in \mathcal{D}(\hat{Q}), \phi(y) > 0, \forall y \in \{0, 1\} \} =
\]

\[
-\inf \{ \lambda p(\frac{\phi(2)}{\phi(1)} - 1) + \mu(1-p)(\frac{\phi(1)}{\phi(2)} - 1); \phi(1), \phi(2) > 0 \}
\]

for \(m = (p, 1-p)\).

The infimum is attained at \(\sqrt{\frac{\mu}{\lambda}(\frac{1}{p} - 1)}\) and \(I(m) = \lambda p + \mu(1-p) - 2\sqrt{\lambda \mu p(1-p)}\).

3.3 Large deviations for stochastic additive functionals

Let us consider the family of stochastic additive functionals \(\xi^\varepsilon(t), t \geq 0\) represented by

\[
\xi^\varepsilon(t) = \xi^\varepsilon(0) + \int_0^t \eta^\varepsilon(ds; x^\varepsilon(\frac{s}{\varepsilon})) , \quad t \geq 0, \varepsilon > 0.
\]

The family of coupled Markov processes \((\xi^\varepsilon(t), x^\varepsilon(\frac{t}{\varepsilon}))\), \(t \geq 0, \varepsilon > 0\) on \(\mathbb{R}^d \times E\) has infinitesimal generator \(\mathbb{L}^\varepsilon\) given by \(\mathbb{L}^\varepsilon = \frac{1}{\varepsilon}Q + Q_1 + \mathbb{I}^\varepsilon(x)\) with the domain \(\mathcal{D}(\mathbb{L}^\varepsilon)\) dense in \(\mathcal{C}(\mathbb{R}^d \times E)\) and the limit process \((\hat{\xi}(t), \hat{x}(t)), t \geq 0\) is a Markov process on \(\mathbb{R}^d \times \hat{E}\).

Our goal is to show the large deviation principle for this family of stochastic additive functionals with the rate function \(I\) stated as

\[
-\inf_{\Gamma_0} I \leq \liminf_{\varepsilon \to 0} \varepsilon \log \mathbb{P}\{\xi^\varepsilon \in \Gamma\} \leq \limsup_{\varepsilon} \varepsilon \log \mathbb{P}\{\xi^\varepsilon \in \Gamma\} \leq -\inf_{\Gamma} I \quad (3.3.1)
\]
where $\Gamma^\circ$ and $\bar{\Gamma}$ represent the interior respectively the closure of the set $\Gamma$. In the particular case in which we take $\Gamma = \{ \xi(t) : ||\xi(t) - \hat{\xi}(t)|| > \delta \}$ one gets the asymptotic behavior of the $\mathbb{P}(\sup_{t \in [0,T]} ||\xi^\varepsilon(t) - \hat{\xi}(t)|| > \delta)$.

**Proposition 3.15** ([13]) If the sequence $\xi^\varepsilon$ satisfies the large deviation principle on $D([0,T], \mathbb{R}^d)$ with rate function $I_u(\varphi)$, then for all bounded continuous functions $h : D([0,T], \mathbb{R}^d) \to \mathbb{R}$,

$$\lim_{\varepsilon \to 0} \varepsilon \log \mathbb{E}\{\exp \left[ -\frac{1}{\varepsilon} h(\xi^\varepsilon) \right] \} = -\inf_{\varphi \in D([0,T], \mathbb{R}^d)} \{ h(\varphi) + I_u(\varphi) \} \quad (3.3.2)$$

The Laplace principle implies the large deviation principle with the same rate function (Theorem 1.2.3).

**Proposition 3.16** If $I_u$ is a rate function on $D([0,T], \mathbb{R}^d)$ and the limit (3.3.2) is valid for all bounded continuous functions $h$, then the sequence $\xi^\varepsilon$ satisfies the large deviation principle on $D([0,T], \mathbb{R}^d)$ with rate function $I$.

**Lemma 3.17** Suppose that for each fixed $k \in \hat{E}$, the family $\xi^\varepsilon_t := \xi^\varepsilon_t(u; k)$, $t \geq 0$, $\varepsilon > 0$ satisfies the large deviation principle with the rate function $I_{u,k}(\cdot)$. If $\hat{x}_t$ is a stationary process on $\hat{E}$ then $\xi^\varepsilon_t(u; \hat{x}(t))$ satisfies the large deviation principle with the rate function $I_u(\varphi) = \min\{I_{u,k}(\varphi) : 1 \leq k \leq N\}$.

**Proof:** Since for each fixed $k \in \hat{E}$, the family $\xi^\varepsilon_t(u; k)$, $t \geq 0$, $\varepsilon > 0$ satisfy the large deviation principle with the rate function $I_{u,k}$, we have

$$\varepsilon \log \mathbb{P}(\xi^\varepsilon \in \Gamma | \hat{x}^\varepsilon = k) \sim -\inf_{\Gamma} I_{u,k}.$$ 

Let’s denote $b^\varepsilon_k := \mathbb{P}(\xi^\varepsilon \in \Gamma | \hat{x}^\varepsilon = k)$, $b_k := \inf_{\Gamma} I_{u,k}$ and $p_k = \mathbb{P}(\hat{x}^\varepsilon = k)$. Thus $\varepsilon \log b^\varepsilon_k \sim -b_k$ and therefore $b^\varepsilon_k = \exp(-\frac{1}{\varepsilon} b_k + c^\varepsilon_k)$ with $c^\varepsilon_k = o(\frac{1}{\varepsilon})$.

We want to prove that $\varepsilon \log \mathbb{P}(\xi^\varepsilon \in \Gamma) \sim -\min\{b_1, \cdots, b_N\}$. We may assume that
\[ b_1 \leq b_2 \leq \cdots \leq b_N \text{ and } 0 < p_i < 1, 1 \leq i \leq N \text{ without loss of generality.} \]

Since \( \mathbb{P}(\xi^e \in \Gamma) = \sum_{k=1}^N \mathbb{P}(\xi^e \in \Gamma|\hat{x}^e = k)\mathbb{P}(\hat{x}^e = k) \), it is enough to prove that 
\[ \varepsilon \log(b_1^e p_1 + \cdots + b_N^e p_N) \sim -b_1 \]
which is equivalent to 
\[ \frac{1}{b_1} \sim \frac{1}{b_i p_1} \]
This is true because 
\[ \frac{b_i}{b_1} = \exp(-\frac{1}{\varepsilon}(b_i - b_1 + \varepsilon(c_i - c_1))) \]
goes to 0 as \( \varepsilon \) goes to 0.

**Theorem 3.18** For absolutely continuous functions \( \varphi \) from \( D([0,T], \mathbb{R}^d) \), with \( T > 0 \) arbitrary fixed, satisfying \( \varphi(0) = u \), and for each fixed \( k \in \hat{E} \), define

\[
I_{u,k}(\varphi) := \int_0^T L(\varphi(t), \dot{\varphi}(t); k) dt,
\]
where \( L \) is subsequently defined by (3.3.9). For all other functions in \( D([0,T], \mathbb{R}^d) \), \( I_{u,k}(\varphi) := \infty \). Then the family \( \xi^e(t), \varepsilon > 0 \) satisfies the large deviation principle with rate function

\[
I_u(\varphi) = \min\{I_{u,k}(\varphi) : 1 \leq k \leq N\}
\]

**Proof:** This will be carried out in several steps. For the sake of clarity, it became necessary to state a number of known results, which we reformulated and adapted to our situation.

**step 1.**

Consider the martingale problem for the generator \( I \times^e \) and its relationship with the exponential martingale problem [8] by taking the transformation \( H^e \) defined as

\[
H^e f := \varepsilon e^{-\frac{1}{\varepsilon}f} \mathbb{L} f e^{-\frac{1}{\varepsilon}f}
\]

An important step is to prove the convergence of \( H^e \) for an appropriate collection of sequences \( f^e \) to an operator \( H \) in the sense that if \( f^e \) converges to \( f \) as \( \varepsilon \to 0 \) the \( H^e f^e \) converges to \( Hf \) [23].
Let us consider the test functions $f^\varepsilon(u, x) = f(u) + \varepsilon \log \varphi^\varepsilon(u, x)$ with $\varphi^\varepsilon(u, x) = \varphi(u, m(x)) + \varepsilon \varphi_1(u, x)$, where $f, \varphi^\varepsilon(u, x)$ are bounded, measurable, continuous differentiable functions on $u \in \mathbb{R}^d$, with bounded first derivative, and uniformly continuous on $E$, convergent to the function $f(u)$. Then, $H^\varepsilon f^\varepsilon$ converges to $H f$,

$$
H f(u; x) := a(u; x) f'(u) + \int_{\mathbb{R}^d} (e^{v f(u)} - 1 - v f'(u)) \Gamma(u, dv; x).
$$

(3.3.6)

Applying the stationary projector $\Pi : \mathcal{B}(E) \to \hat{E}$, defined by

$$
\Pi \varphi(x) := \int_E \rho(dx) \varphi(y) \mathbb{I}(x)
$$

(where $\mathbb{I}(x) = 1$ for all $x \in E$), we obtain

$$
\hat{H} f(u; k) = \hat{a}(u; k) f'(u) + \int_{\mathbb{R}^d} (e^{v f'(u)} - 1 - v f'(u)) \hat{\Gamma}(u, dv; k)
$$

(3.3.7)

where

$$
\hat{a}(u; k) = \int_{E_k} \pi_k(dx) a(u; x) \quad \text{and} \quad \hat{\Gamma}(u, dv; k) = \int_{E_k} \pi_k(dx) \Gamma(u, dv; k).
$$

A key role is played by the function in $u$ and $p$ in $\mathbb{R}^d$ defined by

$$
H(u, p; k) := \hat{a}(u; k) p + \int_{\mathbb{R}^d} (e^{vp} - 1 - vp) \hat{\Gamma}(u, dv; k)
$$

(3.3.8)

having the following properties:

(Ia) for each $p \in \mathbb{R}^d$ and each $k \in \hat{E}$, $\sup_{u \in \mathbb{R}^d} H(u, p; k) < \infty$;

(Ib) for each $k \in \hat{E}$, $H(u, p; k)$ is a continuous function of $(u, p) \in \mathbb{R}^d \times \mathbb{R}^d$. 
For \( u \) and \( q \) in \( \mathbb{R}^d \) we define the Fenchel-Legendre transform

\[
L(u, q; k) := \sup_{p \in \mathbb{R}^d} \{pq - H(u, p; k)\}
\]  

(3.3.9)

**step 2.** As in Lemma 6.2.3. [13] the following properties of the Fenchel-Legendre function can be proved.

**Lemma 3.19** The functions \( H(u, p; k) \) and \( L(u, q; k) \) defined by (3.3.8) and (3.3.9) respectively, have the following properties

(a) For each \( u \in \mathbb{R}^d, k \in \hat{E}, \) \( H(u, p; k) \) is a finite convex function of \( p \in \mathbb{R}^d \) which is differentiable for all \( p \). In addition, \( H(u, p; k) \) is a continuous function of \( (u, p) \in \mathbb{R}^d \times \mathbb{R}^d \)

(b) For each \( u \in \mathbb{R}^d, k \in \hat{E}, \) \( L(u, q; k) \) is a convex function of \( q \in \mathbb{R}^d \). In addition, \( L(u, q; k) \) is a nonnegative, lower semi-continuous function of \( (u, q) \in \mathbb{R}^d \times \mathbb{R}^d \)

(c) \( L(u, q; k) \) is uniformly superlinear in the sense:

\[
\lim_{N \to \infty} \inf_{u \in \mathbb{R}^d} \inf_{q \in \mathbb{R}^d: ||q||=N} \frac{1}{||q||}L(u, q; k) = \infty
\]

(d) For each \( u \in \mathbb{R}^d, k \in \hat{E}, \) the relative interior \( \text{ri}(\text{dom}L(u, \cdot; k)) = \text{ri}(\text{conv}S_{\mu(\cdot|u, k)}) \); in particular \( L(u, q; k) \) equals \( \infty \) for \( u \in \mathbb{R}^d \) and \( q \in (\text{cl}(\text{conv}S_{\mu(\cdot|u, k)}))^c \). For any \( q \in \text{ri}(\text{dom}L(u, \cdot; k)) \) there exists \( v = v(u, q; k) \in \mathbb{R}^d \) such that \( \nabla_u H(u, v(u, q; k); k) = q \). In addition,

\[
L(u, q; k) = v(u, q; k)q - H(u, v(u, q; k); k)
\]
(e) Suppose in addition that for a given \( u \in \mathbb{R}^d \), \( \text{conv}S_{\mu(\cdot|u)} \) has nonempty interior. Then \( H(u,v;k) \) is a strictly convex function of \( v \in \mathbb{R}^d \), \( \text{int}(\text{dom}L(u,\cdot;k)) \) is nonempty, for each \( q \in \text{int}(\text{dom}L(u,\cdot;k)) \) there exists a unique value of \( v \) such that \( \nabla_vH(u,v(u,q;k);k) = q \), and \( L(u,\cdot;k) \) is differentiable on \( \text{int}(\text{dom}L(u,\cdot;k)) \).

(f) For each \( u \) and \( q \) in \( \mathbb{R}^d \), \( k \in \hat{E} \),
\[
L(u,q;k) = \inf\{R(\nu(||\cdot|u,k) : \nu \in \mathcal{P}(\mathbb{R}^d), \int_{\mathbb{R}^d} \nu(dv) = q \}
\]
and the infimum is always attained. If \( L(u,q;k) < \infty \), then the infimum is attained uniquely. \( R(\cdot||\cdot) \) is the relative entropy defined by \( R(\nu||\theta) := \int (\log \frac{d\nu}{d\theta})d\nu \) whenever \( \nu \) is absolutely continuous with respect to \( \theta \). Otherwise \( R(\nu||\theta) := \infty \).

(g) There is a stochastic kernel \( \nu(dv|u,k) \) on \( \mathbb{R}^d \) given \( \mathbb{R}^d \times \hat{E} \) satisfying for \( u \) and \( q \) in \( \mathbb{R}^d \),
\[
R(\nu(\cdot|u,k)||\mu(\cdot|u,k)) = L(u,q;k) \quad \text{and} \quad \int_{\mathbb{R}^d} \nu(dv|u,k) = q
\]

(h) If \( \nu \in \mathcal{P}(\mathbb{R}^d) \) satisfies \( R(\nu(\cdot)||\mu(\cdot|u,k)) < \infty \) for \( u \in \mathbb{R}^d \), \( k \in \hat{E} \) then \( \int_{\mathbb{R}^d} ||v||\nu(dv) < \infty \) and
\[
R(\nu(\cdot|u,k)||\mu(\cdot|u,k)) \geq L(u,\int_{\mathbb{R}^d} \nu(dv);k).
\]

**step 3.** To prove Laplace principle for the sequence \( \xi^\varepsilon \) it is sufficient to prove it for a sequence of random walks \( X^n \) constructed below.
Let $h$ be any bounded continuous function mapping $D([0,T], \mathbb{R}^d)$ into $\mathbb{R}$. We prove the Laplace limit (3.3.2) when $\varepsilon \to 0$ along any sequence $\{\varepsilon_n, n \in \mathbb{N}\}$ converging to 0. Let’s fix such a sequence. By sampling the process $\xi^{\varepsilon_n}$ at a sequence of times depending on $\varepsilon_n$, we define a sequence of piecewise linear processes $\{\zeta^n, n \in \mathbb{N}\}$ for which we prove Laplace principle. Then we show that the sequence is superexponentially closed to $\{\xi^{\varepsilon_n}, n \in \mathbb{N}\}$. Fix $T > 0$. For each $n \in \mathbb{N}$, let $c_n := \left[\frac{T}{\varepsilon_n}\right]$ (where $[x]$ represents the integer part of $x$). Consider the sampled sequence $\xi^{\varepsilon_n}(\frac{Tj}{c_n}), j = 0, 1, \ldots, c_n - 1$. Define $\zeta^n := \{\xi^n(t), t \in [0,T]\}$ by

$$
\zeta^n(t) = \xi^{\varepsilon_n}(\frac{Tj}{c_n}) + c_n(t - \frac{Tj}{c_n}) \left(\xi^{\varepsilon_n}(\frac{T(j+1)}{c_n}) - \xi^{\varepsilon_n}(\frac{Tj}{c_n})\right)
$$

for $t \in \left[\frac{Tj}{c_n}, \frac{T(j+1)}{c_n}\right]$, which is the linear interpolation of the sampled sequence $\xi^{\varepsilon_n}(\frac{Tj}{c_n}), j = 0, 1 \ldots, c_n - 1$.

For each fixed $k \in \hat{E}$, let $\{v^n_j(u; k), u \in \mathbb{R}^d, j \in \mathbb{N}_0\}$ be an i.i.d sequence of random vector fields having the common distribution

$$
\mu^n(dv|u, k) := \mathbb{P}_u \left\{ \frac{c_n}{T} \left( \xi^{\varepsilon_n}(\frac{T}{c_n}) - u \right) \in dv \right\}
$$

(3.3.10)

which is a stochastic kernel on $\mathbb{R}^d$ given $\mathbb{R}^d \times \hat{E}$. We construct the random walks corresponding to the sequence of stochastic kernels $\mu^n(dv|u, k)$ as follows: for each $u \in \mathbb{R}^d, k \in \hat{E}, n \in \mathbb{N}$, consider the sequence of random variables $\{X^n_j, j = 0, 1, \ldots, c_n - 1\}$ taking values in $\mathbb{R}^d$ with

$$
X^n_{j+1} := X^n_j + \frac{T}{c_n} v^n_j(X^n_j; k), \quad X^n_0 = u.
$$
Suppose that the sequence of random vectors $X^n_j$ is interpolated into a piecewise linear continuous-time process $X^n := \{X^n(t), t \in [0, T]\}$ by

$$X^n(t) = X^n_j + \left(t - \frac{Tj}{c_n}\right) v^n_j(X^n_j; k), \quad t \in \left[\frac{Tj}{c_n}, \frac{T(j+1)}{c_n}\right], j = 0, 1, \ldots, c_n - 1$$

Then the distribution of $\zeta^n$ is the same as the distribution of $X^n$. For each $n \in \mathbb{N}$ and $u, p \in \mathbb{R}^d, k \in \hat{E}$, define

$$H^n(u, p; k) := \log \int_{\mathbb{R}^d} e^{vp} \mu^n(dv|u, k)$$

(3.3.11)

**step 4.** We will show that the function $H(u, p; k)$ defined in (3.3.8) can be written as the moment generating function of a stochastic kernel $\mu(dv|u, k)$.

**Lemma 3.20** For $\varepsilon > 0$, $p \in \mathbb{R}^d$, $k \in E$, $s \in [0, T]$, $t \in [s, T]$, $n \in \mathbb{N}$, $u \in \mathbb{R}^d$ and $\delta > 0$, define:

$$N^\varepsilon_p(s; t) := \exp \left[ \frac{1}{\varepsilon} \left( p(\xi^\varepsilon(t) - \xi^\varepsilon(s)) - \int_s^t H(\xi^\varepsilon(v), p; k)dv \right) \right]$$

$$N^\varepsilon_n(t) := \exp \left[ \frac{T}{c_n} \left( p(\xi^n(t) - u) - \int_0^t H(\xi^n(v), p; k)dv \right) \right]$$

and

$$\tau^n_{u, \delta} := \inf \left\{ t \in [0, T] : \inf_{v \in [0, t]} d(\xi^n(v), (B(u, \delta))^c) = 0 \right\}$$

The following conclusions hold:

(a) $N^\varepsilon_p(s; t)$ is an $\mathcal{F}_t^\varepsilon$-martingale for $t \in [s, T]$ and therefore

$$\mathbb{E}_\mu \{ N^\varepsilon_p(s; t) \} = \mathbb{E}_\mu \{ N^\varepsilon_p(s; s) \} = 1$$
(b) $N^\varepsilon_p(0; t \wedge \tau^n_{u,\delta})$ is an $\mathcal{F}_n^\varepsilon$-martingale for $t \in [0, T]$, hence
\[ \mathbb{E}_u\{N^\varepsilon_p(0; t \wedge \tau^n_{u,\delta})\} = \mathbb{E}_u\{N^\varepsilon_p(0; 0)\} = 1 \]

(c) For any $p \in \mathbb{R}^d$,
\[ \sup_n \sup_{u \in \mathbb{R}^d} \mathbb{E}_u \left\{ \left( \frac{T}{\sqrt{c_n}} \right) \right\} \leq 1 \]
and
\[ \sup_n \sup_{u \in \mathbb{R}^d} \mathbb{E}_u \left\{ \left( \frac{T}{\sqrt{\tau^n_{u,\delta}}} \right) \right\} \leq 1 \]

(d) There exists $\gamma_1 \in (0, \infty)$ such that for any $u, p \in \mathbb{R}^d$
\[ H(u, p) \geq \hat{p} \hat{u}(u; k) \geq -\gamma_1 ||p|| \]

(e) For any $p \in \mathbb{R}^d$, there exists $\gamma_2 \in \mathbb{R}$ such that
\[ \sup_n \sup_{u \in \mathbb{R}^d} \mathbb{E}_u \left\{ \left( \frac{T}{\sqrt{c_n}} \right)^2 \right\} \leq e^{\gamma_2} \]
and
\[ \sup_n \sup_{u \in \mathbb{R}^d} \mathbb{E}_u \left\{ \left( \frac{T}{\sqrt{\tau^n_{u,\delta}}} \right)^2 \right\} \leq e^{\gamma_2} \]

(f) For any $p \in \mathbb{R}^d$ and compact subset $K \subseteq \mathbb{R}^d$, there exists $M \in \mathbb{N}$ such that
\[ \inf_n \inf_{u \in K} \mathbb{E}_u \left\{ \left( \frac{T}{\sqrt{c_n}} \wedge \tau^n_{u,\delta} \right) \right\} \geq e^{-\gamma_3(\delta)}, \]
where $\gamma_3(\delta) \to 0$ as $\delta \to 0$.

Proof:

(a) Part (a) follows from Th 4.2.1 [8]

(b) Since $\tau^n_{u,\delta}$ is a stopping time, using the optional stopping theorem, part (b) follows: Proposition 2.1.5, Theorem 2.2.13 [9]. Characterization and convergence.
(c) Define \( p_n := \frac{T}{\xi_n \varepsilon_n} \in [1, \infty) \), then \( \left[ N_{p,u}^\varepsilon_n(t) \right]^{p_n} = N_p^\varepsilon_n(0;t) \). Using Hölder inequalities we get

\[
\sup_{n \in \mathbb{N}} \sup_{u \in \mathbb{R}^d} \mathbb{E}_u \left\{ \sim N_{p,u}^\varepsilon_n(t) \right\} \leq \sup_{n \in \mathbb{N}} \sup_{u \in \mathbb{R}^d} \left( \mathbb{E}_u \left\{ N_p^\varepsilon_n(0;t) \right\} \right)^{\frac{1}{p_n}} = 1
\]

\[
\sup_{n \in \mathbb{N}} \sup_{u \in \mathbb{R}^d} \mathbb{E}_u \left\{ \sim N_{p,u}^\varepsilon_n \left( \frac{T}{c_n} \wedge \tau_{u,\delta}^n \right) \right\} \leq \sup_{n \in \mathbb{N}} \sup_{u \in \mathbb{R}^d} \left( \mathbb{E}_u \left\{ N_p^\varepsilon_n(0;t \wedge \tau_{u,\delta}^n) \right\} \right)^{\frac{1}{p_n}} = 1
\]

(d) Since \( e^z - 1 - z \geq 0 \forall z \in \mathbb{R} \), \( H(u, p; k) \geq p \hat{a}(u; k) \forall u, p \in \mathbb{R}^d \). Since \( \hat{a}(u; k) \) is bounded it follows that there exists \( \gamma_1 \in (0, \infty) \) such that \( p \hat{a}(u; k) \geq -\gamma_1 ||p|| \).

(e) In order to prove the first inequality, we write

\[
\sim N_{2p,u}^\varepsilon_n \left( \frac{T}{c_n} \right) \exp \left[ c_n \frac{T}{T} \int_0^{\frac{T}{c_n} \wedge \tau_{u,\delta}^n} H(\xi^\varepsilon_n(v), 2p; k) dv - 2 c_n \frac{T}{T} \int_0^{\frac{T}{c_n} \wedge \tau_{u,\delta}^n} H(\xi^\varepsilon_n(v), 2p; k) dv \right]
\]

\[
\leq \sim N_{2p,u}^\varepsilon_n \left( \frac{T}{c_n} \right) \exp(\gamma_2) \text{ where }
\gamma_2 := \sup_{u \in \mathbb{R}^d} H(u, 2p; k) - 2 \inf_{u \in \mathbb{R}^d} H(u, p; k)
\]

which is finite because of the condition (Ia) and part (d) of this lemma.

Similarly for the second inequality.

(f) For \( n \in \mathbb{N} \), define \( \bar{p}_n := c_n \frac{\xi_n}{T - \varepsilon_n} \geq 1. \)

\[
\Gamma_{u,\delta}^n := c_n \frac{T}{T - \varepsilon_n} \int_0^{\frac{T}{c_n} \wedge \tau_{u,\delta}^n} H(\xi^\varepsilon_n(v), 2p; k) dv - c_n \frac{T}{T - \varepsilon_n} \int_0^{\frac{T}{c_n} \wedge \tau_{u,\delta}^n} H(\xi^\varepsilon_n(v), T(T - \varepsilon_n)p; k) dv
\]
For any $u \in \mathbb{R}^d$, part(b) and Hölder inequality imply:

$$1 = \left( \mathbb{E}_u \left\{ N_{T,\epsilon_n}^{\tau_n}(0; \frac{T}{c_n}, \tau_{n,\epsilon_n}) \right\} \right)^{\bar{p}_n} \leq \mathbb{E}_u \left\{ N_{T,\epsilon_n}^{\tau_n}(0; \frac{T}{c_n}, \tau_{n,\epsilon_n}) \right\}^{\bar{p}_n}$$

$$= \mathbb{E}_u \left\{ \sim_{\epsilon_n} N_{p,u}^{\tau_n}(\frac{T}{c_n}, \tau_{n,\epsilon_n}) \exp(\Gamma_{n,\epsilon_n}) \right\}.$$

Since $\epsilon_n \to 0$, there exists $M \in \mathbb{N}$ such that $\epsilon_n \leq \frac{1}{2} \wedge \delta$ for any $n \geq M$. Hence for any compact subset $K \subset \mathbb{R}^d$,

$$\sup_{n \geq M} \sup_{u \in K} \sup |H(v, p; k) - \frac{1}{T - \epsilon_n} H(v, T(T - \epsilon_n); p; k)| \leq \gamma_3(\delta)$$

where $\gamma_3(\delta)$ equals

$$3 \sup_{u \in K} \sup |H(y, q; k) - H(u, p; k)| : ||y - u|| \leq \delta, ||q - p|| \leq \delta + 2\delta \sup_{u \in K} |H(u, p; k)|$$

Thus

$$\inf_{n \in \mathbb{N}} \inf_{u \in K} \mathbb{E}_u \left\{ \sim_{\epsilon_n} N_{p,u}^{\tau_n}(\frac{T}{c_n}, \tau_{n,\epsilon_n}) \right\} \geq e^{-\gamma_3(\delta)}$$

Condition (Ia) and part (d) of this lemma imply that $\sup_{u \in K} |H(u, p; k)| < \infty$ and the uniform continuity of $H$ on compact subsets of $\mathbb{R}^d \times \mathbb{R}^d$ gives us $\gamma_3(\delta) \to 0$ as $\delta \to 0$.

Since the conditions of the Proposition 10.3.2 in [13] are fulfilled the next result follow.

**Proposition 3.21** For each $k \in \hat{E}$, the following conclusions hold:
(a) there exists a superlinear function \( f : (0, \infty) \to \mathbb{R} \cup \{\infty\} \) such that for any \( \varepsilon > 0, \delta > 0, s \in [0, T], t \in (s, T] \)

\[
\sup_{u \in \mathbb{R}^d} \mathbb{P}_{u,k} \left\{ \sup_{s \leq \sigma \leq t} ||\xi^\varepsilon(\sigma) - \xi^\varepsilon(s)|| \geq \delta \right\} \leq 2d \exp \left( -\frac{t-s}{\varepsilon} f(\frac{\delta}{\sqrt{d}(t-s)}) \right)
\]

(b) for each \( p \in \mathbb{R}^d, \sup_{u \in \mathbb{R}^d} \sup_{u \in \mathbb{R}^d} H^n(u, p; k) < \infty \)

(c) for each \( p \in \mathbb{R}^d \) and each compact subset \( K \subset \mathbb{R}^d \),

\[
\lim_{n \to \infty} \sup_{u \in K} |H^n(u, p; k) - H(u, p; k)| = 0 \quad (3.3.12)
\]

(d) for each \( u \in \mathbb{R}^d \), the sequence of probability measures \( \mu^n(dv|u, k), n \in \mathbb{N} \) converges weakly to a probability measure \( \mu(dv|u, k) \) on \( \mathbb{R}^d \) and for each \( p \in \mathbb{R}^d \),

\[
H(u, p; k) = \log \int_{\mathbb{R}^d} e^{p u} \mu(dv|u, k).
\]

The family \( \mu(dv|u, k), u \in \mathbb{R}^d, k \in \hat{E} \) defines a stochastic kernel on \( \mathbb{R}^d \) given \( \mathbb{R}^d \times \hat{E} \). In addition, the function mapping \( u \in \mathbb{R}^d \mapsto \mu(\cdot|u, k) \in \mathcal{P}(\mathbb{R}^d) \) is continuous in the topology of weak convergence on \( \mathcal{P}(\mathbb{R}^d) \).

**Definition 3.22** A measurable function \( f : (0, \infty) \to \mathbb{R} \cup \{\infty\} \) is called superlinear if \( \lim_{c \to \infty} \frac{f(c)}{c} = \infty \).

**Proof:**

(a) For \( i \in \{1, \ldots, d\} \) let \( u^i \) be the unit vector in the \( i^{th} \) coordinate direction of \( \mathbb{R}^d \) and let \( r > 0 \) be a real number. Then for any \( \varepsilon > 0, \delta > 0, s \in [0, T], t \in (s, T] \), we have:

\[
\sup_{u \in \mathbb{R}^d} \mathbb{P}_{u} \left\{ \sup_{s \leq \sigma \leq t} < \pm u^i, \xi^\varepsilon(\sigma) - \xi^\varepsilon(s) > \geq \frac{\delta}{\sqrt{d}} \right\} =
\]

\[
\sup_{u \in \mathbb{R}^d} \mathbb{P}_{u} \left\{ \sup_{s \leq \sigma \leq t} < \pm ru^i, \xi^\varepsilon(\sigma) - \xi^\varepsilon(s) > \geq \frac{r\delta}{\sqrt{d}} \right\} \leq
\]
\[
\sup_{u \in \mathbb{R}^d} \mathbb{P}_u \left\{ \sup_{s \leq \sigma \leq t} \exp \left( \frac{1}{\varepsilon} < \pm ru^i, \xi^\varepsilon(\sigma) - \xi^\varepsilon(s) > - \int_s^\sigma H(\xi^\varepsilon(\sigma), \pm ru^i; k) d\sigma \right) \geq \exp \left( \frac{1}{\varepsilon} \left( \frac{r\delta}{\sqrt{d}} - (t - s)\bar{H}(r; k) \right) \right) \right\}
\]

where \( \bar{H}(r; k) := \max_{\pm} \max_{i=1,...,d} \sup_{u \in \mathbb{R}^d} H(u, \pm ru^i; k) \). The property (Ib) implies that \( \bar{H} \) is finite for each \( r > 0 \).

For \( c > 0 \) define \( f(c) := \sup_{r > 0} \{ rc - \bar{H}(r) \} \). Then for \( M > 0 \) and all sufficiently large \( c \), \( \frac{1}{c} f(c) \geq M - \frac{1}{c} \bar{H}(M) \geq \frac{M}{2} \). Since \( M \) is arbitrary chosen it implies that \( f \) is superlinear. We rewrite the above inequality in terms of \( N^\varepsilon_{\pm ru^i}(s; \sigma) \) and then use the submartingale inequality. We get

\[
\sup_{u \in \mathbb{R}^d} \mathbb{P} \left\{ \sup_{s \leq \sigma \leq t} \frac{1}{\varepsilon} < \pm u^i, \xi^\varepsilon(\sigma) - \xi^\varepsilon(s) > \geq \frac{\delta}{\sqrt{d}} \right\} \leq \sup_{u \in \mathbb{R}^d} \mathbb{P}_u \left\{ \sup_{s \leq \sigma \leq t} N^\varepsilon_{\pm ru^i}(s; \sigma) \geq \exp \left[ \frac{1}{\varepsilon} \left( \frac{r\delta}{\sqrt{d}} - (t - s)\bar{H}(r) \right) \right] \right\} \leq \exp \left[ - \frac{1}{\varepsilon} \left( \frac{r\delta}{\sqrt{d}} - (t - s)\bar{H}(r) \right) \right] \sup_{u \in \mathbb{R}^d} \mathbb{P}_u \left\{ N^\varepsilon_{\pm ru^i}(s; t) \right\} = \exp \left[ - \frac{1}{\varepsilon} \left( \frac{r\delta}{\sqrt{d}} - (t - s)\bar{H}(r) \right) \right] \leq \exp \left[ - \frac{t - s}{\varepsilon} \frac{r\delta}{\sqrt{d}(t - s)} \bar{H}(r) \right] = \exp \left[ - \frac{t - s}{\varepsilon} f \left( \frac{\delta}{\sqrt{d}(t - s)} \right) \right].
\]

Combining these bounds for the \( 2d \) choices of vectors \( \{ \pm u^i, i = 1, ..., d \} \), we get

\[
\sup_{u \in \mathbb{R}^d} \mathbb{P}_u \left\{ \sup_{s \leq \sigma \leq t} || \xi^\varepsilon(\sigma) - \xi^\varepsilon(s) || \geq \delta \right\} \leq 2d \exp \left[ - \frac{t - s}{\varepsilon} f \left( \frac{\delta}{\sqrt{d}(t - s)} \right) \right].
\]
\(\exp\{H^n(u, p; k)\} = \exp\left(\log \mathbb{E}_u \left\{ \exp \left( \frac{c_n}{T} \left( \xi^{\varepsilon_n}(T) - u \right) \right) \right\} \right)\)

\[= \mathbb{E}_u \left\{ \exp \left( \frac{c_n}{T} \left( \xi^{\varepsilon_n}(T) - u \right) \right) \right\}\]

\[= \mathbb{E}_u \left\{ \sim^\varepsilon_n \mathcal{N}_{u, p}(\frac{T}{c_n}) \exp \left[ \frac{c_n}{T} \int_0^T H(\xi^{\varepsilon_n}(v), p; k) dv \right] \right\}\]

Therefore,

\[\sup_{n \in \mathbb{N}} \sup_{u \in \mathbb{R}^d} \exp\{H^n(u, p; k)\} \leq \sup_{n \in \mathbb{N}} \sup_{u \in \mathbb{R}^d} \mathbb{E}_u \left\{ \sim^\varepsilon_n \mathcal{N}_{u, p}(\frac{T}{c_n}) \exp \left( \sup_{v \in \mathbb{R}^d} H(v, p; k) \right) \right\}\]

\[\leq \exp \left( \sup_{v \in \mathbb{R}^d} H(v, p; k) \right) < \infty\]

(c) According to (Ib), for any \(p \in \mathbb{R}^d\), \(H(u, p; k)\) is a uniformly continuous function on \(u\) in any compact subset of \(\mathbb{R}^d\). Thus, given \(\theta > 0\), there exists \(\delta > 0\) such that

\[\sup_{u \in \mathcal{K}} \sup_{\{v: ||v - u|| \leq \delta\}} |H(v, p; k) - H(u, p; k)| < \theta\]

For \(n \in \mathbb{N}\), define: \(A(\varepsilon_n, u, \delta) := \left\{ \sup_{0 \leq s \leq \frac{T}{c_n}} ||\xi^{\varepsilon_n}(s) - u|| < \delta \right\}\). Then using

(a) we have

\[\sup_{u \in \mathbb{R}^d} \mathbb{P}_u \{A(\varepsilon_n, u, \delta)^c\} \leq 2d \exp \left( -\frac{T}{c_n \varepsilon_n} f\left(\frac{\delta c_n}{\sqrt{dT}}\right) \right)\]

\[\leq 2d \exp \left( -f\left(\frac{\delta c_n}{\sqrt{dT}}\right) \right)\]
For $u \in K, p \in \mathbb{R}^d$,

$$
\exp[H^n(u, p; k) - H(u, p; k)] \\
= \mathbb{E}_u \left\{ \exp \left( \frac{c_n}{T} \left( \xi^{(n)} \left( \frac{T}{c_n} \right) - u \right) \right) - H(u, p; k) \right\} \\
= \mathbb{E}_u \left\{ \sim_{n} \mathcal{N}_{u,p} \left( \frac{T}{c_n} \right) \exp \left[ \frac{c_n}{T} \int_{0}^{\frac{c_n}{T} \wedge \tau^n_{u,\delta}} (H(\xi^{(n)}(t), p; k) - H(u, p; k)) dt \right] \right\}.
$$

From the definition of $\tau^n_{u,\delta}$ and part (f) of lemma it follows that there exists $M \in \mathbb{N}$ such that

$$
\inf_{n \geq M} \inf_{u \in K} \mathbb{E}_u \left\{ \sim_{n} \mathcal{N}_{u,p} \left( \frac{T}{c_n} \wedge \tau^n_{u,\delta} \right) \exp \left[ \frac{c_n}{T} \int_{0}^{\frac{c_n}{T} \wedge \tau^n_{u,\delta}} (H(\xi^{(n)}(t), p; k) - H(u, p; k)) dt \right] \right\} \\
\geq \exp(-\gamma_3(\delta) - \theta) \text{ with } \gamma_3(\delta) \to 0 \text{ as } \delta \to 0, \text{ and,}
$$

$$
\sup_{n \geq M} \sup_{u \in K} \mathbb{E}_u \left\{ \sim_{n} \mathcal{N}_{u,p} \left( \frac{T}{c_n} \wedge \tau^n_{u,\delta} \right) \exp \left[ \frac{c_n}{T} \int_{0}^{\frac{c_n}{T} \wedge \tau^n_{u,\delta}} (H(\xi^{(n)}(t), p; k) - H(u, p; k)) dt \right] \right\} \\
\leq \exp(\theta) \text{ (from part (c) of lemma and the definition of } \tau^n_{u,\delta}).
$$

Let $\gamma_4 = \sup_{v \in \mathbb{R}^d} H(v, p; k) - \inf_{u \in K} H(u, p; k) < \infty$ and $A(\varepsilon_n, u, \delta) = \{\tau^n_{u,\delta} > \frac{T}{c_n}\}$. Then

$$
\sup_{u \in K} \left| \mathbb{E}_u \left\{ \sim_{n} \mathcal{N}_{p,u} \left( \frac{T}{c_n} \right) \exp \left[ \frac{c_n}{T} \int_{0}^{\frac{T}{c_n} \wedge \tau^n_{u,\delta}} (H(\xi^{(n)}(t), p; k) - H(u, p; k)) dt \right] \right\} \right| - \\
\mathbb{E}_u \left\{ \sim_{n} \mathcal{N}_{p,u} \left( \frac{T}{c_n} \wedge \tau^n_{u,\delta} \right) \exp \left[ \frac{c_n}{T} \int_{0}^{\frac{T}{c_n} \wedge \tau^n_{u,\delta}} (H(\xi^{(n)}(t), p; k) - H(u, p; k)) dt \right] \right\} \exp(\gamma_4) \sup_{u \in K} \mathbb{E}_u \left\{ \mathbb{1}_{(A(\varepsilon_n, u, \delta))} \mathcal{N}_{p,u} \left( \frac{T}{c_n} \right) \right\} \leq \\
\mathbb{E}_u \left\{ \mathbb{1}_{(A(\varepsilon_n, u, \delta))} \mathcal{N}_{p,u} \left( \frac{T}{c_n} \wedge \tau^n_{u,\delta} \right) \right\}
$$
\[ \leq \exp(\gamma_4)(\sup_{u \in K}(E_u(I_{A(x_n,u,\delta)}c))^{1/2}(E_u(N_{u,p}(\frac{T}{c_n})^2))^{1/2} \]
\[ + \sup_{u \in K}(E_u(I_{A(x_n,u,\delta)}c))^{1/2}(E_u(N_{u,p}(\frac{T}{c_n})^2))^{1/2} \leq \]
\[ 2\sqrt{2d}\exp\left(\gamma_4 - f\left(\frac{\delta c_n}{T\sqrt{d}}\right)\right)\exp\left(\frac{\gamma_2}{2}\right) \]

which goes to 0 as \( n \to \infty \).

Therefore
\[ \lim \inf \inf_{n \to \infty} \sup_{u \in K} [H^n(u, p; k) - H(u, p; k)] \geq -\gamma_3(\delta) - \theta \]
\[ \lim \sup \sup_{n \to \infty} [H^n(u, p; k) - H(u, p; k)] \leq \theta \]
so
\[ \lim \sup_{n \to \infty} \sup_{u \in K} |H^n(u, p; k) - H(u, p; k)| = 0. \]

(d) Condition (Ib) and lemma (c) implies that for each \( u \in \mathbb{R}^d \) and \( k \in E \), the moment generating function of \( \{\mu^n(\cdot/u, k), n \in \mathbb{N}\} \) (which is the sequence \( \{\exp(H^n(u, p; k))\} \)), converges for each \( p \in \mathbb{R}^d \) to \( \exp(H(u, p; k)) \). Since the latter is a continuous function, we conclude using the continuity of moment generating functions that \( \mu^n(\cdot/u, k) \Rightarrow \mu(\cdot/u, k) \) where \( \mu(\cdot/u, k) \) is a probability measure on \( \mathbb{R}^d \) satisfying for each \( p \in \mathbb{R}^d \)
\[ H(u, p; k) = \log \int_{\mathbb{R}^d} \exp(pv)\mu(dv/u, k). \]

Thus the family \( \{\mu(\cdot/u, k), u \in \mathbb{R}^d\} \) is a stochastic kernel on \( \mathbb{R}^d \) (it is a limit of a stochastic kernel). The continuity theorem for moment generating function and (Ib) implies the continuity of the map \( u \in \mathbb{R}^d \to \mu(\cdot/u, k) \).

\[ \blacksquare \]
step 5.

In order to study the Laplace principle for the process $X^n$, we need to verify the asymptotic behavior of

$$W^n(u) := -\frac{1}{c_n} \log \mathbb{E}_u(\exp(-c_n h(X^n))),$$  \hspace{1cm} (3.3.13)

where $\mathbb{E}_u$ denotes the expectation with respect to $\mathbb{P}_u$ and $h$ is any bounded continuous function mapping $C([0,T], \mathbb{R}^d)$ into $\mathbb{R}$. We will show that this is equal to the minimal cost of function of an associated stochastic control problem.

We now specify the stochastic control problem whose minimal cost function gives a representation for the function $W^n(u)$. The controlled process is a discrete-time process $\bar{X}_j^n$, $j = 0, 1, \cdots, c_n - 1$, and at each time $t$ there will be a control $\nu^n_j$ giving the distributions of the controlled random variable that replaces this noise due to the increments. $\nu^n_j$ is a stochastic kernels on $(\mathbb{R}^d)^{j+1}$, denoted by $\nu^n_j(dv) = \nu^n_j(dv|\bar{X}_j^n, \cdots, \bar{X}_0^n)$. A sequence of controls $\{\nu^n_{1,j}, j = 0, 1, \cdots, c_n - 1\}$ is called an admissible control sequence.

Then, as in [13] (Theorem 4.3.1) we get the variational representation of $W^n_u$ as

$$W^n(u) = \inf_{\nu^n_j} \mathbb{E}_u \left\{ \sum_{j=0}^{c_n-1} \left[ \frac{1}{c_n} R(\nu^n_j(\cdot)||\mu(\cdot|\bar{X}_j^n, k)) \right] + h(\bar{X}_j^n) \right\}$$  \hspace{1cm} (3.3.14)

where the infimum is taken over all admissible control sequences $\{\nu^n_j\}$. For $n \in \mathbb{N}$ and $t \in [0,T]$, define the stochastic kernel

$$\nu^n(dv|t) := \begin{cases} 
\nu^n_j(dv), & t \in \left( \frac{T_j}{c_n}, \frac{T(j+1)}{c_n} \right), \\
\nu^n_{c_n-1}(dv), & t \in \left[ \frac{T(c_n-1)}{c_n}, T \right] 
\end{cases} \hspace{1cm} j = 0, 1, \cdots, c_n - 2$$
The following representation holds (similar as in [13] (Corollary 5.2.1))

\[ W^n(u) = \inf_{\nu^n_1} \mathbb{E}_u \left\{ \int_0^T R(\nu^n_1(\cdot|t)||\mu(\cdot|\tilde{X}^n(t)) + h(\tilde{X}^n) \right\} \]  \hspace{1cm} (3.3.15)

where \( \tilde{X}^n = \{ \tilde{X}^n(t), t \in [0, T] \} \) is the piecewise constant interpolation of the controlled random variables \( \{ \tilde{X}^n_j, j = 0, 1, \cdots, c_n - 1 \} \).

**step 6.** Laplace principle upper bound

Let \( I_{u,k}(\varphi) := \int_0^T L(\varphi(t), \dot{\varphi}(t), k) dt \) where \( L \) is the Legendre-Fenchel transform defined in (3.3.9). Then \( I_{u,k} \) is a rate function and

\[ \limsup_{n \to \infty} \frac{1}{c_n} \log \mathbb{E}_u(\exp(-c_n h(X^n))) \leq -\inf_{\varphi \in C([0,T],\mathbb{R}^d)} (I_{u,k}(\varphi) + h(\varphi)) \]  \hspace{1cm} (3.3.16)

Indeed, first it can be shown that \( I_{u,k} \) has compact level sets in \( C([0,T],\mathbb{R}^d) \) by using parts (b) and (c) of the Proposition 3.21, which implies that \( I_{u,k} \) is a rate function. Then using part (h) of Proposition 3.19 we will get

\[ \liminf_{n \to \infty} W^n(u) \geq \inf_{\varphi \in C([0,T],\mathbb{R}^d)} (I_{u,k}(\varphi) + h(\varphi)). \]

**step 7.** Laplace principle lower bound.

In order to prove the Laplace principle lower bound we need to characterize the relative interior of the effective domain of \( L(u, \cdot|k) \) in terms of the stochastic kernel \( \mu(\cdot|u,k) \). This is done in part (d) of Proposition 3.19.

For \( A, B \) subsets of \( \mathbb{R}^d \) define

\[ A + B := \{ u \in \mathbb{R}^d : u = a + b, a \in A, b \in B \}. \]
A subset $C$ of $\mathbb{R}^d$ is called a convex cone if it has the property that for $c \in C$, $\lambda c \in C$ \( \forall \lambda \in [0, \infty) \). Denote $conC$ for the convex cone of $C$.

We can rewrite $H(u, p; k)$ as

$$H(u, p; k) = \hat{b}(u; k)p + \int_{\mathbb{R}^d} (e^{xp} - 1)\hat{\Gamma}(u, dv; k)$$

where

$$\hat{b}(u; k) := \hat{a}(u; k) - \int_{\mathbb{R}^d} v\hat{\Gamma}(u, dv; k)$$

Let $S_{\hat{\Gamma}(u, k)}$ be the support of $\hat{\Gamma}(u, k)$ and define $T_{(u, k)} := \{\hat{b}(u; k)\} + conS_{\hat{\Gamma}(u, k)}$.

The relative interior $ri(domL(u, \cdot; k)) = ri(T_{(u, k)})$ and the following properties hold:

(a) The sets $intT_{(u, k)}$ are independent of $(u, k) \in \mathbb{R}^d \times \hat{E}$

(b) $0 \in intT_{(u, k)}$

With similar arguments as in Theorem 6.5.1 [13] it can be proved that

$$\limsup_{n \to \infty} W_n(u, k) \leq \inf_{\varphi \in C([0, T], \mathbb{R}^d)} (I_{uk}(\varphi) + h(\varphi)).$$

This gives the Laplace principle lower bound for $X^n$.

$$\liminf_{n \to \infty} \frac{1}{c_n} \log \mathbb{E}_u(\exp(-c_n h(X^n))) \geq - \inf_{\varphi \in C([0, T], \mathbb{R}^d)} (I_{uk}(\varphi) + h(\varphi)). \quad (3.3.17)$$

Thus the Laplace principle is proved for the random walk $X^n$ and therefore for the process $\zeta^n$. 
**step 8.** Laplace principle holds for the sequence $\xi^n_t$ because $\xi^n_t, \zeta^n_t$ are super-exponentially closed, i.e.

$$\limsup_{n \to \infty} \sup_{u \in \mathbb{R}^d} \varepsilon_n \log \mathbb{P}_u (\rho(\xi^n_t, \zeta^n_t) > \delta) = -\infty,$$

where $\rho$ is Skorokhod metric on $D([0, T], \mathbb{R}^d)$.

Thus, by Proposition 3.16 we obtain the large deviation principle for the sequence of random variables $\xi^n_t(u; k)$ with the rate function

$$I_{u,k}(\varphi) = \int_0^T L(\varphi(t), \dot{\varphi}(t); k)dt.$$

Using Lemma 3.17 we get the large deviation principle for the sequence of stochastic additive functionals $\xi^n_t$ with rate function $I_u(\varphi) = \min \{ I_{u,k}(\varphi) : 1 \leq k \leq N \}$.

This completes the proof of the theorem.

This principle has many applications, for example finding the probability of exit from a stable domain of the process. In some cases the infimum can be explicitly found by using calculus of variations. The class of absolutely continuous functions on $[0, T]$ can be identified with the Sobolev space $H^{1,1}[0, T]$, and since the Fenchel-Legendre function $L(u, q; k)$ verifies the conditions of Tonelli’s existence theorem (Theorem 3.7 [24]), the existence of the minimizer will follow. If $\varphi \in AC[0, T]$ is a local minimizer of the functional $L(\varphi, \varphi')$, then $\varphi$ will satisfy the Euler-Lagrange equation which will be further simplified to the Beltrami equation: $L(\varphi, \varphi') - \varphi' L_{\varphi'}(\varphi, \varphi') = C$, where $C$ is a constant.

**Example 3.23 Compound Poisson process**
Consider the compound Poisson process $\xi^\epsilon(t), t \geq 0$ switched by the jump Markov process $x(t), t \geq 0$ defined in Example 3.13, of the form

$$
\xi^\epsilon(t) = \sum_{k=1}^{\nu(t/\epsilon;x(t/\epsilon))} a_k(x(\frac{t}{\epsilon}))
$$

with the infinitesimal generator given by

$$
\Pi^\epsilon(x)\phi(u) = \frac{\Lambda(x)}{\epsilon} \int_{\mathbb{R}^d} [\phi(u + \epsilon v) - \phi(u)] F(dv; x).
$$

Here $\nu(t; x), t \geq 0, x \in E = \{1, 2, 3, 4\}$ is a homogeneous Poisson process, with intensity $\Lambda(x)$ and $a_k(x), k \geq 1, x \in E$ is a sequence of i.i.d. random variables, independent of $\nu(t), t \geq 0$, with common distribution $F(dv; x)$.

Using notation $\hat{a}(k) = \int_{E_k} \pi_k(dx)a(x)$, this process converges weakly,

$$
\xi^\epsilon(t) \Rightarrow \int_0^t \hat{a}(\hat{x}(s))ds, \quad \text{as } \epsilon \to 0.
$$

Applying the operator $H^\epsilon f^\epsilon$ as in equation (3.3.5) we get the limiting operator $Hf$ as follows

$$
Hf(u, x) = \Lambda(x) \int_{\mathbb{R}^d} [e^{vf(u)} - 1] F(dv, x).
$$

For tractability purposes, let’s suppose that $F(dv; x)$ is independent of $x$. Then the projected operator $\hat{H}f$ is

$$
\hat{H}f(u, k) = \hat{\Lambda}(k) \int_{\mathbb{R}^d} [e^{vf(u)} - 1] F(dv)
$$

where $\hat{\Lambda}(k) = \int_{E_k} \pi_k(dx)\Lambda(x)$. Hence, $\hat{\Lambda}(1) = \frac{2\lambda_1\mu_1}{\lambda_1 + \mu_1}$ and $\hat{\Lambda}(2) = \frac{2\lambda_2\mu_2}{\lambda_2 + \mu_2}$. Assume that the random variables $a_k(x)$ are distributed exponential with the parameter $\lambda$. Then
the function $H(p; k), p \in \mathbb{R}, k \in \hat{E} = \{1, 2\}$ defined in the relation 3.3.8 is

$$H(p; k) = \hat{\Lambda}(k) \frac{p}{\lambda - p}, \quad \lambda > p$$

The Legendre-Fenchel transform $L(q; k) = \sup_{p \in \mathbb{R}} \{pq - H(p; k)\}$ becomes

$$L(q; k) = \lambda q - 2\sqrt{\lambda q \hat{\Lambda}(k)} + \hat{\Lambda}(k),$$

the supremum being attained for $p = \lambda - \sqrt{\frac{\lambda \hat{\Lambda}(k)}{q}}$. Therefore, for $T > 0$ arbitrary fixed, and for absolutely continuous functions $\varphi \in D([0, T], \mathbb{R})$, with $\varphi(0) = 0$, the process $\xi$ satisfies the large deviation principle. Its rate function is $I(\varphi) = \min_{k=1,2} I_k(\varphi)$, where

$$I_k(\varphi) = \int_0^T L(\varphi'(t); k) dt$$

and

$$L(\varphi'(t)) = \lambda \varphi'(t) - 2\sqrt{\lambda \varphi'(t) \hat{\Lambda}(k)} + \hat{\Lambda}(k).$$
CHAPTER 4
LARGE DEVIATION PRINCIPLE FOR FUNCTIONAL CENTRAL LIMIT THEOREMS

We study a class of empirical measures on $\mathcal{C}[0, \infty)$ associated with $S_n(g) = \sum_{k=0}^{n-1} g(X_k)$, where $(X_k)$ is a Markov chain with an invariant measure $m$ and $g \in L^2(m)$, via an invariance principle corresponding to interpolation of $S_n$.

The necessary background is introduced and auxiliary results are provided for subsequent analysis. Then we provide criteria which allow us to establish a martingale decomposition representation for $S_n(g)$, the key element in the proof of the large deviation result presented in the last section of this chapter.

4.1 Functional a.e. central limit theorems for additive functionals

Let $X_1, X_2, \ldots$ be a sequence of i.i.d. random variables on $(\Omega, \mathcal{F}, \mathbb{P})$ such that $\mathbb{E}(X_i) = 0$, $\mathbb{E}(X_i^2) = 1$ and $S_n = \sum_{i=1}^{n} X_i$. Define the interpolation processes, with $
abla_n(\frac{k}{n}) = S_k$ for $1 \leq k \leq n$,

$$\nabla_n(t) := \frac{1}{\sqrt{n}} \left( S_{\lfloor nt \rfloor} + (nt - \lfloor nt \rfloor)(S_{\lfloor nt \rfloor + 1} - S_{\lfloor nt \rfloor}) \right), \quad 0 \leq t \leq 1$$

(4.1.1)

and empirical processes $W_n : \mathcal{C}[0, 1] \to \mathcal{M}_1(\mathcal{C}[0, 1])$,

$$W_n(\cdot) := \frac{1}{L(n)} \sum_{k=1}^{n} \frac{1}{k} \delta_{(\nabla_k \in \cdot)}$$

(4.1.2)

with $L(n) = \sum_{k=1}^{n} \frac{1}{k}$. The following functional almost everywhere central limit theorem is due to Brosamler [25].
Theorem 4.1  The random process $W_n$ converges weakly to the Wiener measure $W$ on $C[0,1]$, $\mathbb{P}$-a.e.

Let $(X_n)$ be an ergodic Markov chain with stationary measure $m$, $S_n(g) = \sum_{k=0}^{n-1} g(X_k)$ be such that $\mathbb{E}_m(S_1^2) = \sigma^2 \in (0,\infty)$ and for any initial distribution $\mu$ and any $k \in \mathbb{N}$, $\mathbb{E}_\mu(S_k^2) < \infty$. Define the interpolation processes

$$\Psi_n(t) := \frac{1}{\sigma \sqrt{n}} \left( S_{[nt]} + (nt-[nt])(S_{[nt]+1} - S_{[nt]}) \right), \quad 0 \leq t < \infty$$

(4.1.3)

and the empirical processes

$$W_n(\cdot) := \frac{1}{L(n)} \sum_{k=1}^{n} \frac{1}{k} \delta_{\{\Psi_k \in \cdot\}}$$

(4.1.4)

Theorem 4.2  $W_n$ converges weakly to the Wiener measure $W$ on $C[0,\infty)$, $\mathbb{P}_\mu$-a.e.

Proof: By using our martingale decomposition (see section 3, Theorem 4.11) we can write $S_n = M_n + R_n$, where $(M_n)$ is a mean zero martingale for which $\sup_{1 \leq k \leq n} \frac{|R_k|}{\sigma \sqrt{n}}$ converges in probability to 0. Define the empirical measures

$$W_n^M(\cdot) = \frac{1}{L(n)} \sum_{k=1}^{n} \frac{1}{k} \delta_{\{\Psi_k^M \in \cdot\}}$$

(4.1.5)

where $\Psi_n^M$ is the interpolation process

$$\Psi_n^M(t) = \frac{1}{\sigma \sqrt{n}} \{M_{[nt]} + (nt-[nt])(M_{[nt]+1} - M_{[nt]}) \right)$$

(4.1.6)
corresponding to the martingale $M = (M_n)$. By [26], $W^M_n$ converges weakly to the Wiener measure $W$ on $C[0, 1]$, i.e., for any bounded continuous function $f : C[0, 1] \rightarrow \mathbb{R}$, $W$- a.e.,

$$\lim_{n \to \infty} \frac{1}{L(n)} \sum_{k=1}^{n} \frac{1}{k} \delta_{\Psi_k^M} = \phi(W)$$

To show that $W_n$ converges weakly to the Wiener measure $W$ on $C[0, 1]$, we check that

$$\lim_{n \to \infty} \frac{1}{L(n)} \sum_{k=1}^{n} \frac{1}{k} \delta_{\Psi_k^M} = \lim_{n \to \infty} \frac{1}{L(n)} \sum_{k=1}^{n} \frac{1}{k} \delta_{\Psi_k^M},$$

for which it is sufficient to have $\lim_{n \to \infty} ||\Psi_n - \Psi_n^M||_{C[0, 1]} = 0$. Indeed,

$$||\Psi_n - \Psi_n^M||_{C[0, 1]} = \sup_{t \in [0, 1]} |\Psi_n(t) - \Psi_n^M(t)| =$$

$$\sup_{t \in [0, 1]} \left| \frac{1}{\sigma \sqrt{n}} \left\{ R_{[nt]} + (nt - [nt])(R_{[nt] + 1} - R_{[nt]}) \right\} \right| \leq$$

$$\sup_{1 \leq k \leq n} \sup_{t \in \left[ \frac{k-1}{n}, \frac{k}{n} \right]} \frac{1}{\sigma \sqrt{n}} |R_{[nt]} + (nt - [nt])(R_{[nt] + 1} - R_{[nt]})| \leq$$

$$\sup_{1 \leq k \leq n} \frac{1}{\sigma \sqrt{n}} |R_k|,$$

which by (4.5.2) converges to 0 in probability.

By replacing 1 with $N$ in the above proof the result holds for $C[0, N]$, for every $N > 0$. The extension to $C[0, \infty)$ is standard: $\nu_n$ converges weakly to $\nu$ on $C[0, \infty)$ if for every $N > 0$, $\pi_N \nu_n$ converges weakly to $\pi_N \nu$ on $C[0, N]$, where $\pi_N$ of a measure on $C[0, \infty)$ denotes the measure it induces on $C[0, N]$.

In our case this means that $W_n$ converges weakly to the Wiener measure $W$ on $C[0, \infty)$. 

$\blacksquare$
4.2 Large deviation principle for Ornstein-Uhlenbeck processes

Define the canonical shifts $\sigma_t : C(\mathbb{R}) \to C(\mathbb{R})$ by

$$\sigma_t \omega(s) := \omega(t+s)$$

and empirical processes $R_T : C(\mathbb{R}) \to \mathcal{M}_1(C(\mathbb{R}))$ for any $T > 0$ by

$$R_T := \frac{1}{2T} \int_{-T}^{T} \delta_{\sigma_t \omega} dt.$$ 

Assume that $\sigma$ is an ergodic transformation, and let $\mathbb{P}$ be a $\sigma$-invariant measure i.e. $\mathbb{P}$ is the probability measure on the space of trajectories corresponding to the Ornstein-Uhlenbeck processes starting with stationary distribution at time $t = 0$.

Let $\Omega = C(\mathbb{R})$, then for any bounded continuous function we have:

$$\int_{\Omega} f(\omega) dR_T(\omega) = \frac{1}{2T} \int_{-T}^{T} \left( \int_{\Omega} f(\omega) d\delta_{\sigma_t \omega} \right) dt = \frac{1}{2T} \int_{-T}^{T} f(\sigma_t \omega) dt 
\to \int_{\Omega} f(\omega) d\mathbb{P}$$

a.s. as $T \to \infty$ by ergodic theorem. This implies the weak convergence $R_T \Rightarrow \mathbb{P}$ as $T \to \infty$. The exponential decay for the deviations of $R_T$ from $\mathbb{P}$ was given by Donsker and Varadhan in [27] through the rate function $H(Q), Q \in \mathcal{M}_1(C(\mathbb{R}))$ defined as

$$H(Q) := \begin{cases}
\lim_{T \to \infty} \frac{1}{T} h(\mathbb{E}_{[-T:T]}(Q)/\mathbb{E}_{[-T:T]}(\mathbb{P})) & \text{if } Q \text{ is } \sigma \text{ invariant} \\
\infty & \text{otherwise}
\end{cases}$$
where $\Im|_I(\mu)$ denote the image of a measure $\mu$ under the restriction map on $I$ and $h(\mu/\nu)$ is the relative entropy of $\mu$ with respect to $\nu$,

$$h(\mu/\nu) := \begin{cases} 
\int \log \left( \frac{d\mu}{d\nu} \right) & \text{if } \mu \ll \nu \\
\infty & \text{otherwise}
\end{cases} \quad \text{(4.2.1)}$$

**Theorem 4.3** For every bounded interval $I \subset \mathbb{R}$,

$$H_I(\cdot) := \inf_{Q \in \Im|_I^{-1}(\cdot)} H(Q)$$

is a rate function on $\mathcal{M}_1(\mathcal{C}(I))$ and for any Borel set $A \subseteq \mathcal{M}_1(\mathcal{C}(I))$,

$$-\inf_A H_I \leq \liminf_{T \to \infty} \frac{1}{T} \log \mathbb{P}\{\Im|_I(R_T) \in A\} \leq \limsup_{T \to \infty} \frac{1}{T} \log \mathbb{P}\{\Im|_I(R_T) \in A\} \leq -\inf_A H_I$$

An extension to paths on $[0, \infty)$ by using a finer topology than the topology of uniform convergence on compact sets was obtained by Heck in [28].

**Theorem 4.4** Let $\Phi : \mathbb{R} \to [0, \infty)$ be a continuous function satisfying

$$\lim_{t \to \infty} \frac{\Phi(t)}{\sqrt{t}} = \lim_{t \to -\infty} \frac{\Phi(t)}{\sqrt{|t|}} = \infty \quad \text{(4.2.2)}$$

and $\mathcal{C}_\Phi := \{\omega \in \mathcal{C}(\mathbb{R}) : \sup_{t \in \mathbb{R}} \frac{|\omega(t)|}{\Phi(t)} < \infty\}$, $\mathcal{M}_1(\mathcal{C}_\Phi) = \{Q \in \mathcal{M}_1(\mathcal{C}(\mathbb{R})) : Q(\mathcal{C}_\Phi) = 1\}$.

Then $H|_{\mathcal{M}_1(\mathcal{C}_\Phi)}$ is a rate function on $\mathcal{M}_1(\mathcal{C}_\Phi)$, and for every Borel set $A \subseteq \mathcal{M}_1(\mathcal{C}_\Phi)$,

$$-\inf_A H \leq \liminf_{T \to \infty} \frac{1}{T} \log \mathbb{P}\{R_T \in A\} \leq \limsup_{T \to \infty} \frac{1}{T} \log \mathbb{P}\{R_T \in A\} \leq -\inf_A H$$
Remark 4.5 **Under condition (4.2.2), \( P \in \mathcal{M}_1(\mathcal{C}_\phi) \) and \( P \)-a.e. \( R_T \in \mathcal{M}_1(\mathcal{C}_\phi) \).**

### 4.3 Large deviation principle for Brownian motion

Let \( \phi : \mathbb{R}_+ \to \mathbb{R}_+ \) be a continuous function such that

\[
\lim_{t \to 0} \frac{\phi(t)}{\sqrt{t \log t}} = \lim_{t \to \infty} \frac{\phi(t)}{\sqrt{t \log t}} = \infty, \tag{4.3.1}
\]

and the set \( \mathcal{C}_\phi \) defined as

\[
\mathcal{C}_\phi := \{ \omega \in \mathcal{C}[0, \infty) : \sup_{t \in \mathbb{R}_+} \frac{|\omega(t)|}{\phi(t)} < \infty \}
\]

Consider the isomorphism \( F : \mathcal{C}([0, \infty)) \to \mathcal{C}([0, \infty)) \), \( F(\omega)(t) := \sqrt{t} \omega(\log t) \). Then \( F|_{\mathcal{C}_\phi} : \mathcal{C}_\phi \to \mathcal{C}_\phi \) is a bijective isometry with respect to \( \cdot \) and \( \cdot \) if \( \Phi(\cdot) = \frac{\Phi(\exp(\cdot))}{\sqrt{\exp(\cdot)}} \).

Let us denote \( W := \mathcal{Z}_F(P) \), so \( W \) is Wiener measure on \( \mathcal{C}[0, \infty) \). Theorem 4.4 can be written in terms of Brownian motion as follows:

For any \( t \in \mathbb{R} \) define \( \theta_t : \mathcal{C}[0, \infty) \to \mathcal{C}[0, \infty) \) by \( \theta_t \omega(s) := e^{-t/2} \omega(e^t s) \) and for any \( T > 0 \) empirical processes \( S_T : \mathcal{C}[0, \infty) \to \mathcal{M}_1(\mathcal{C}[0, \infty)) \) by

\[
S_T(\omega) := \frac{1}{2T} \int_{-T}^{T} \delta_{\theta_t \omega} dt
\]

**Theorem 4.6** Define for any \( Q \) in \( \mathcal{M}_1(\mathcal{C}[0, \infty)) \) the function

\[
I(Q) := \begin{cases} 
\lim_{T \to \infty} \frac{1}{T} h(\mathbb{S} |_{[e^{-T}, e^T]}(Q) / \mathbb{S} |_{[e^{-T}, e^T]}(W)) & \text{if } Q \text{ is } \theta \text{-invariant} \\
\infty & \text{otherwise}
\end{cases} \tag{4.3.2}
\]

Axiom 4.7 **If \( P \) satisfies condition (4.2.2) and \( \mathcal{C}_\phi \) is a compact subset of \( \mathcal{C}_\phi \) then \( \mathcal{C}_\phi \) is a compact subset of \( \mathcal{C}_\phi \).**

**Theorem 4.8** Define for any \( Q \) in \( \mathcal{M}_1(\mathcal{C}[0, \infty)) \) the function

\[
I(Q) := \begin{cases} 
\lim_{T \to \infty} \frac{1}{T} h(\mathbb{S} |_{[e^{-T}, e^T]}(Q) / \mathbb{S} |_{[e^{-T}, e^T]}(W)) & \text{if } Q \text{ is } \theta \text{-invariant} \\
\infty & \text{otherwise}
\end{cases} \tag{4.3.2}
\]
Then $I|_{\mathcal{M}_1(C_\phi)}$ is a rate function and for any Borel set $A \subseteq \mathcal{M}_1(C_\phi)$,

$$-\inf_A I \leq \liminf_{T \to \infty} \frac{1}{T} \log W\{S_T \in A\} \leq \limsup_{T \to \infty} \frac{1}{T} \log W\{S_T \in A\} \leq -\inf_A I$$

where $\mathcal{M}_1(C_\phi) := \{Q \in \mathcal{M}_1(C[0,\infty)) : Q(C_\phi) = 1\}$

Let’s define for any $a > 0$, $\theta_a : C[0,\infty) \to C[0,\infty)$ by

$$\theta_a \omega(t) = \frac{1}{\sqrt{a}} \omega(at)$$

and empirical processes $R_T : C[0,\infty) \to \mathcal{M}_1(C[0,\infty))$,

$$R_T := \frac{1}{\log T} \int_1^{\sqrt{T}} \delta_{\theta_t \omega} \frac{dt}{t}$$

Then $(R_n)$ satisfies LDP with constants $(\log n)$ and rate function $I$, where

$$I(Q) := \begin{cases} 
\lim_{a \to \infty} \frac{1}{2 \log a} h \left( Q \circ |\frac{1}{a},\frac{1}{a}| / W \circ |\frac{1}{a},\frac{1}{a}| \right) & \text{if } Q \text{ is } \theta\text{-invariant} \\
\infty & \text{otherwise}
\end{cases}$$

which means

$$-\inf_A I \leq \liminf_{n \to \infty} \frac{1}{\log n} \log W\{R_n \in A\} \leq \limsup_{n \to \infty} \frac{1}{\log n} \log W\{R_n \in A\} \leq -\inf_A I$$

Since $\frac{1}{\log n} \int_1^n \delta_{\theta_t \omega} \frac{dt}{T}$ has under $W$ the same distribution as $\frac{1}{\log n} \int_{\frac{1}{\sqrt{a}}}^{\sqrt{n}} \delta_{\theta_t \omega} \frac{dt}{T}$ we have the following corollary:
Corollary 4.7 Let $\tilde{R}_n = \frac{1}{\log n} \int_1^n \delta_t dt$. Then $(\tilde{R}_n)$ satisfies LDP with constants $(\log n)$ and rate function $I$ defined in (4.3.3).

4.4 Large deviation principle for sequence of i.i.d. random variables

Let’s consider $X_1, X_2, \ldots$ i.i.d. random variables on $(\Omega, \mathcal{F}, \mathbb{P})$ such that $\mathbb{E}(X_i) = 0$, $\mathbb{E}(X_i^2) = 1$ and $S_n = \sum_{i=1}^n X_i$.

We want to prove the LDP for the functional almost everywhere central limit theorem given in Theorem 4.1.

Lemma 4.8 Let $Y_n$ and $Z_n$ be random variables with values in a metric space $(E, d)$ such that for all $\epsilon > 0$,

$$\lim_{n \to \infty} \frac{1}{\log n} \log \mathbb{P}\{d(Y_n, Z_n) > \epsilon\} = -\infty$$

(4.4.1)

Then $Y_n$ and $Z_n$ are equivalent with respect to LDP (or exponentially equivalent), which means that if $(Y_n)$ satisfies LDP with constants $(\log n)$ and rate function $I$, then $(Z_n)$ also satisfies LDP with the same constants and rate function.

To use this lemma for sequences of random variables in $\mathcal{M}_1(C_\phi)$ it is necessary to define a metric, as done in [28].

Let $(C_\phi, | \cdot |_\phi)$ be a metric space, with $|\omega\rangle_\phi = \sup_{t \in \mathbb{R}^+} \frac{|\omega(t)|}{\phi(t)}$. On $\mathcal{M}_1(C_\phi)$ define

$$d_\phi(\mu, \nu) := \sup \left\{ \left| \int f d\mu - \int f d\nu \right|, f \in C(C_\phi, \mathbb{R}), \|f\|_L \leq \frac{1}{2} \right\}$$

(4.4.2)

where $\|f\|_L := \sup_{\omega \in C_\phi} |f(\omega)| + \sup_{\omega, \omega' \in C_\phi, \omega \neq \omega'} \frac{|f(\omega) - f(\omega')|}{|\omega - \omega'|_\phi}$. Then $(\mathcal{M}_1(C_\phi), d_\phi)$ becomes a metric space and the following properties hold:

(a) $d_\phi \leq 1$
(b) \( d_\phi(\alpha \mu + (1 - \alpha) \nu, \alpha \tilde{\mu} + (1 - \alpha) \tilde{\nu}) \leq \alpha d_\phi(\mu, \tilde{\mu}) + (1 - \alpha) d_\phi(\nu, \tilde{\nu}), \) for \( \alpha \in [0, 1] \)
and \( \mu, \tilde{\mu}, \nu, \tilde{\nu} \in M_1(C_\phi) \)
\( (c) \ d_\phi(\delta_\omega, \delta_\omega') \leq |\omega - \omega'|_{\psi} \) for \( \omega, \omega' \in C_\phi \).

Then, as shown in [29], we have

Lemma 4.9 Let \( Y_t, Z_t, t \in [1, \infty) \) be random variables with values in the separable metric space \( (C_\phi, | \cdot |_\phi) \) such that \( t \to Y_t, t \to Z_t \) are continuous from the right or left and for all \( \varepsilon > 0 \),

\[
\lim_{t \to \infty} \frac{1}{\log t} \log \mathbb{P} \left\{ \sup_{s \in [t, t+1)} |Y_s - Z_s|_{\phi} > \varepsilon \right\} = -\infty \quad (4.4.3)
\]

Then \( \left( \frac{1}{\log n} \int_1^n \delta_{Y_s} \frac{ds}{s} \right) \) and \( \left( \frac{1}{\log n} \int_1^n \delta_{Z_s} \frac{ds}{s} \right) \) are equivalent with respect to LDP.

Let \( W_n = \frac{1}{L(n)} \sum_{k=1}^n \frac{1}{k} \delta_{\psi_k} \) be as defined in (4.1.2) and \( \phi \) above satisfies in addition \( \phi(t) > t^{1+\varepsilon} \), for \( t \in [0, 1] \).

Theorem 4.10 \((W_n)\) satisfies LDP with constants \((\log n)\) and rate function \( I \) defined by (4.3.3).

Proof:

Let \( \tilde{W}_n := \frac{1}{\log n} \int_1^n (\delta_{\psi_s}) \frac{ds}{s} \). First we show that \( \tilde{W}_n \) and \( W_n \) are exponentially equivalent. By Lemma 4.8 it suffices to verify

\[
\lim_{n \to \infty} \frac{1}{\log n} \log \mathbb{P} \{ d_\phi(W_n, \tilde{W}_n) > \varepsilon \} = -\infty
\]
Given $f \in C(C_{\phi}, \mathbb{R})$, $\|f\|_L \leq \frac{1}{2}$ we have

$$|\int f \, dW_n - \int f \, d\tilde{W}_n| = \left| \frac{1}{\log n} \int_{\frac{1}{n}}^{n} f(\psi_t) \, dt - \frac{1}{L(n)} \sum_{k=1}^{n} \frac{1}{k} f(\psi_k) \right|$$

$$= \left| \frac{1}{\log n} \sum_{k=1}^{n-1} \log(1 + \frac{1}{k}) \delta_{\psi_k} - \frac{1}{L(n)} \sum_{k=1}^{n} \frac{1}{k} f(\psi_k) \right|$$

$$\leq \left| \frac{1}{L(n)} \log(1 + \frac{1}{n} f(\psi_n)) \right| + \left| \frac{1}{L(n)} \sum_{k=1}^{n-1} \left( \log(1 + \frac{1}{k}) - \frac{1}{k} \right) f(\psi_k) \right|$$

$$+ \left| \frac{1}{L(n)} - \frac{1}{\log n} \right| \left| \sum_{k=1}^{n-1} \log(1 + \frac{1}{k} f(\psi_k)) \right|$$

$$\leq \frac{1}{2} \left( \frac{1}{nL(n)} + \frac{C}{L(n)} \sum_{k=1}^{n-1} \frac{1}{k^2} + \left| 1 - \frac{\log n}{L(n)} \right| \right)$$

which converges to 0 uniformly as $n \to \infty$. Let $X_1, X_2, ...$ be i.i.d. random variables such that $\mathbb{E}(X_i) = 0$, $\mathbb{E}(X_i^2) = 1$ and the partial sum $S_n = \sum_{i=1}^{n} X_i$. If one shows that $(\tilde{W}_n)$ satisfies the LDP with constants $(\log n)$ and rate function $I$ in the case in which $X_i$ are i.i.d. $N(0, 1)$-distributed, then the general case follows, see [29], i.e., by Skorokhod’s representation theorem there exists a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with random variables $(Y_n)$ and $(Z_n)$ such that $Y_n$, $n \in \mathbb{N}$ are i.i.d. with the same distribution as $X_n$, $Z_n$ are i.i.d $N(0, 1)$-distributed and $(\tilde{W}_n)$ corresponding to $(Y_n)$ and respectively $(Z_n)$ are shown to be equivalent with respect to LDP. Therefore, it remains to consider the case for $X_i$ which are i.i.d. $N(0, 1)$-distributed.

Let $(\Omega = \mathcal{C}_0[0, \infty), \mathcal{F}, \mathbb{P} = W)$ with coordinate map $X_t(\omega) = \omega(t)$ and $X_t(\omega) := \omega(i) - \omega(i - 1) \sim$ i.i.d. $N(0, 1)$-distributed. We will show that $(\tilde{W}_n)$ satisfies the LDP by checking that $\tilde{W}_n$ and $\tilde{R}_n$ are equivalent with respect to LDP.
By Lemma 4.9, this reduces to checking that for any \( \epsilon > 0 \)

\[
\lim_{n \to \infty} \frac{1}{\log n} \log W \left\{ \sup_{s \in [n, n+1]} |\theta_s - \psi_n| \phi > \epsilon \right\} = -\infty \tag{4.4.4}
\]

The probability \( W\{\sup_{s \in [n, n+1]} |\theta_s - \psi_n| \phi > \epsilon \} \) is dominated by the sum of three probabilities, \( W\{\sup_{s \in [n, n+1]} |\theta_n - \psi_n| \phi > \epsilon/2 \}, \)

\( W\{\sup_{s \in [n, n+1]} \sup_{t \in \mathbb{R}^+} \frac{1}{\sqrt{s}} |\omega(st) - \omega(nt)|/\phi(t) > \epsilon/4 \}, \)

\( W\{\sup_{s \in [n, n+1]} \sup_{t \in \mathbb{R}^+} |\omega(nt)|/\phi(t) > \epsilon/4 \}. \) Given \( \epsilon > 0 \), based on arguments of Lemma 4 in [28], the estimates below hold for large \( n \) and utilize the properties of function \( \phi \) given in Theorem 4.6 along with the properties of Brownian motion,

\[
W \left\{ \omega \in C_0[0, \infty) : \sup_{t \in [a, b]} |\omega(t) - \omega(a)| \geq c \right\} \leq 2 \exp \left( -\frac{c^2}{2(b-a)} \right)
\]

for all \( 0 \leq a < b \) and \( c > 0 \), and the inequalities: \( \log(e^{|k|}) \geq \frac{|k|+1}{2}, \sup_{s \in [n, n+1]} (\frac{1}{\sqrt{n}} - \frac{1}{\sqrt{n}}) \leq \frac{1}{2} n^{-3/2} \). In what follows the positive constant \( c_{ij} \) represent the \( j \)-th constant in the \( i \)-th estimate.

Regarding the first estimate we have

\[
W \left\{ \sup_{s \in [n, n+1]} \sup_{t \in \mathbb{R}^+} |\theta_s - \psi_n| \phi > \frac{\epsilon}{2} \right\} \leq \\
W \left\{ \sup_{k \geq 1} \sup_{t \in [\frac{k}{n}, \frac{k+1}{n}]} \frac{|\theta_n(t) - \psi_n(t)|}{\phi(t)} > \frac{\epsilon}{2} \sqrt{n} \right\} \leq \\
W \left\{ \sup_{k \geq 1} \sup_{t \in [k, k+1]} |\omega(t) - \omega(k)| > \frac{\epsilon}{2} \sqrt{n} \phi \left( \frac{k}{n} \right) \right\}
\]
Then for sufficiently large \( n \)

\[
\begin{align*}
W \left\{ \sup_{1 \leq k \leq n} \sup_{t \in [k, k+1)} |\omega(t) - \omega(k)| > \frac{\varepsilon}{2} \sqrt{n} \phi \left( \frac{k}{n} \right) \right\} \leq \\
\sum_{k=1}^{n} W \left\{ \sup_{t \in [k, k+1)} |\omega(t) - \omega(k)| > \frac{\varepsilon}{2} \sqrt{n} \frac{k}{n} \left( \frac{k}{n} \right)^{\varepsilon} \right\} \leq \\
\sum_{k=1}^{n} 2 \exp \left( -\frac{1}{2} \frac{\varepsilon^{2}}{4} \frac{k^{1+\varepsilon}}{n^{2\varepsilon}} \right) \leq 2n \exp \left( -\frac{\varepsilon^{2}}{8} \frac{n}{n} \right)
\end{align*}
\]

and

\[
\begin{align*}
W \left\{ \sup_{k \geq n+1} \sup_{t \in [k, k+1)} |\omega(t) - \omega(k)| > \frac{\varepsilon}{2} \sqrt{n} \log \left( \frac{k}{n} \right) \right\} \leq \\
\sum_{k=n+1}^{\infty} W \left\{ \sup_{t \in [k, k+1)} |\omega(t) - \omega(k)| > \frac{\varepsilon}{2} \sqrt{3k} \right\} \leq \\
\sum_{k=n+1}^{\infty} 2 \exp \left( -\frac{1}{2} \frac{\varepsilon^{2}}{4} 9k \right) \leq c_{11} e^{-c_{12}n}.
\end{align*}
\]

Therefore,

\[
W \left\{ \sup_{s \in [n, n+1]} |\theta_s - \psi_{n, \phi}| > \varepsilon \right\} \leq 2n \exp \left( -\frac{\varepsilon^{2}}{8} \frac{n}{n} \right) + c_{11} e^{-c_{12}n}.
\]

For the second estimate one obtains
\[ W \left\{ \sup_{s \in [n,n+1)} \sup_{t \in \mathbb{R}^+} \frac{\omega(st) - \omega(nt)}{\sqrt{s} \phi(t)} > \frac{\varepsilon}{4} \right\} \leq \]

\[ \sum_{l \in \mathbb{Z}} W \left\{ \sup_{s \in [n,n+1)} \sup_{t \in [e^l, e^{l+1})} \frac{|\omega(st) - \omega(nt)|}{\phi(t)} > \frac{\varepsilon}{4} \right\} \leq \]

\[ \sum_{l \in \mathbb{Z}} W \left\{ \sup_{s \in [0,1)} \sup_{t \in [e^l, e^{l+1})} \frac{|\omega(s + nt) - \omega(nt)|}{\phi(t)} > \frac{\varepsilon}{4} \sqrt{n} \phi(e^l) \right\} \leq \]

\[ \sum_{l \in \mathbb{Z}} W \left\{ \sup_{s \in [0,1)} \sup_{t \in [1,e^l)} \frac{|\omega((s + n)t) - \omega(nt)|}{\phi(t)} > \frac{\varepsilon}{4} \sqrt{n \log(e^l / \zeta(e^l))} \right\} \leq \]

\[ \sum_{l \in \mathbb{Z}} c_{21} n \exp(-c_{22} \varepsilon^2 n (|l| + 1)) \leq c_{23} n \exp(-c_{24} \varepsilon^2 n), \]

where \( \zeta : [1, \infty) \rightarrow [1, \infty) \),

\[ \zeta(t) := \inf_{s \in \mathbb{R}^+ \setminus \{(t, t)\}} \frac{\phi(s)^2}{s \log(s)}, \lim_{t \to \infty} \zeta(t) = \infty. \]

Finally,

\[ W \left\{ \sup_{s \in [n,n+1)} \sup_{t \in \mathbb{R}^+} \left| \frac{1}{\sqrt{s}} - \frac{1}{\sqrt{n}} \right| \frac{\omega(nt)}{\phi(t)} > \frac{\varepsilon}{4} \right\} \leq \]

\[ W \left\{ \sup_{t \in \mathbb{R}^+} \frac{\omega(nt)}{n \sqrt{n} \phi(t)} > \frac{\varepsilon}{4} \right\} \leq \sum_{l \in \mathbb{Z}} W \left\{ \sup_{t \in [e^l, e^{l+1})} \frac{\omega(t)}{\phi(t)} > \frac{\varepsilon n}{4} \right\} \leq \]

\[ \sum_{l \in \mathbb{Z}} W \left\{ \sup_{t \in [1,e^l)} |w(t)| > \frac{\varepsilon n}{4} \sqrt{\log(e^l / \zeta(e^l))} \right\} \leq \]

\[ \sum_{l \in \mathbb{Z}} c_{31} \exp(-\varepsilon^2 n^2 c_{32} (|l| + 1)) \leq c_{32} \exp(-c_{34} \varepsilon^2 n^2). \]

which establishes (4.4.4) and concludes the proof.
4.5 Martingale decomposition of Markov functionals

In this section we show that certain functionals of Markov chains can be written as martingales perturbed by random variables whose maxima (scaled by factor $\frac{1}{\sqrt{n}}$) converge in probability to zero at a prescribed rate. The main interest in such representation stems from the fact that when proving central limit theorem or large deviation principle, the methods in Markovian versus martingale case are essentially different (except when the two cases overlap and either method can be employed), while martingale approach is often preferable. For instance, martingale maximal inequalities offer a useful tool in handling the tail estimates which in turn could be applied to Markovian cases whenever possible.

In what follows we provide criteria for a martingale decomposition and describe the classes of Markov chain functionals that allow such representation.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, $(E, \mathcal{B}_E)$ a complete separable metric space, and let $X_n, n \geq 0$ be a Markov chain defined on $\Omega$ with values in $E$. Given a probability measure $\mu$ on $(E, \mathcal{B}_E)$, one defines the probability measures $\mathbb{P}_\mu$ by

$$\mathbb{P}_\mu(B) = \mu \mathbb{P}(B) = \int_E \mu(dx)p(x, B), \quad x \in E, B \in \mathcal{F}$$

where $p(x, B), x \in E, B \in \mathcal{B}_E$ is the Markov transition function of $X_n, n \geq 0$. We denote by $\mathbb{E}_\mu$ and $\mathbb{E}_x$ the expectations corresponding to $\mathbb{P}_\mu$ and $\mathbb{P}_x$ respectively. Inductively one defines the $n$-step transition probability by $p^n(x, B) = \mathbb{P}(X_n \in B | X_0 = x) = \mathbb{P}(X_{n+m} \in B | X_m = x)$.

Let $P$ be the transition probability operator defined as

$$P \varphi(x) := \mathbb{E}[\varphi(x_{n+1}) | x_n = x] = \int_E p(x, dy) \varphi(y)$$
and denote by $P^n$ the $n$-step transition operator corresponding to the $n$-step transition probability $p^n(x,B)$.

**Theorem 4.11** Let $X_n,n \geq 0$ be an ergodic $E$-valued Markov chain with initial distribution $\mu$ and unique invariant probability measure $m$.

If $g \in L^2(m) := \{ g : E \to \mathbb{R} : \int_E g^2 \, dm < \infty \}$ satisfies the properties:

(i) $\int_E g \, dm = 0$

(ii) $\| P^k g \|_{L^2(m)} \leq \rho^k \| g \|_{L^2(m)}$ for some $0 < \rho < 1$, $k \in \mathbb{N}$

(iii) $\frac{d p^k}{d m} \leq D < \infty$, $k \in \mathbb{N}$ and $\int_{\{x : g^2(x) > n\}} g^2(x) m(dx) \leq \exp(-\varphi(n))$ for $n$ large,

with $\varphi : \mathbb{R}_+ \to \mathbb{R}_+$ such that $\lim_{x \to \infty} \frac{\varphi(x)}{\log(x)} = \infty$

(iv) $|P^k g(x)| \leq C n$, whenever $|g(x)| \leq n$, for some $1 < C < \infty$, $x \in E$, $k \in \mathbb{N}$, and large $n$

then

$$S_n(g) := \sum_{k=0}^{n-1} g(X_k) = M_n + R_n \quad (4.5.1)$$

where $M_n$ is a mean zero martingale relative to $(\Omega, \mathcal{F}_n, \mathbb{P})$, with the natural filtration $\mathcal{F}_n = \sigma(X_0, \ldots, X_n)$ and

$$\lim_{n \to \infty} \frac{1}{\log n} \log \mathbb{P}\left\{ \sup_{1 \leq k \leq n} \frac{R_k^2}{n} > \varepsilon \right\} = -\infty. \quad (4.5.2)$$

**Proof:** We show that under our hypothesis the Poisson equation $(I - P)u = g$ has a unique solution. Notice that $I - P$ is not invertible, while $((1+\varepsilon)I - P)u_\varepsilon = g$, $\varepsilon > 0$ has unique solution

$$u_\varepsilon = ((1+\varepsilon)I - P)^{-1} g = \frac{1}{1+\varepsilon} \sum_{k=1}^{\infty} \frac{P^{k-1}g}{(1+\varepsilon)^{k-1}} \quad (4.5.3)$$
thanks to $1 + \varepsilon$ being in the resolvent of $L^2(m)$-contractive $P$, whose spectral radius is 1. Condition (ii) implies that the series in (4.5.3) converges in $L^2(m)$ and we have

\[ u(x) := \lim_{\varepsilon \to 0} u_\varepsilon(x) = \sum_{i=0}^{\infty} P^i g. \]  

Therefore, $S_n(g) = \sum_{k=0}^{n-1} g(X_k) = \sum_{k=1}^{n} (u(X_k) - Pu(X_{k-1})) + u(X_0) - u(X_n)$ and by Markov property

\[ M_n := \sum_{k=1}^{n} (u(X_k) - Pu(X_{k-1})) \]

is a mean zero martingale in $L^2(m)$ with respect to the natural filtration $\mathcal{F}_n = \sigma(X_0, ..., X_n)$. Setting

\[ R_n := u(X_0) - u(X_n) \]

establishes the martingale decomposition (4.5.1).

To simplify notation, summation index below and whenever else integer is needed, is understood as the integer part of the number. For each $k \in \mathbb{N}$ we have

\[
\mathbb{P}(u^2(X_k) > Cn) \leq \mathbb{P} \left( \left| \sum_{i=0}^{\lfloor \sqrt{n} \rfloor} P^i g(X_k) \right| > \frac{\sqrt{Cn}}{2} \right) + \mathbb{P} \left( \left| \sum_{i=\lfloor \sqrt{n} \rfloor + 1}^{\infty} P^i g(X_k) \right| > \frac{\sqrt{Cn}}{2} \right)
\]
The second term with $n \geq 4$ satisfies

$$\mathbb{P} \left( \left| \sum_{i=\lceil \sqrt{n} \rceil+1}^{\infty} P^i g(X_k) \right| > \frac{\sqrt{Cn}}{2} \right) \leq \mathbb{E} \left| \sum_{i=\lceil \sqrt{n} \rceil+1}^{\infty} P^i g(X_k) \right| =$$

$$\mathbb{E} \left( \sum_{i=\lceil \sqrt{n} \rceil+1}^{\infty} P^i g(X_k) \right)_{L^1(\Omega)} \leq \mathbb{E} \left( \sum_{i=\lceil \sqrt{n} \rceil+1}^{\infty} P^i g(X_k) \right)_{L^2(\Omega)}$$

$$\sum_{i=\lceil \sqrt{n} \rceil+1}^{\infty} \|P^i g(X_k)\|_{L^2(\Omega)}$$

Using conditions (iii) and (ii) we get

$$\|P^i g(X_k)\|_{L^2(\Omega)}^2 = \mathbb{E}(P^i g(X_k))^2 = \mathbb{E} \left( \int g(y)p^i(x,dy) \right)^2 = \int \left( \int g(y)p^i(x,dy) \right)^2 \mu^k(dx) \leq D \|P^i g(x)\|_{L^2(\Omega(m))}^2 D \rho^2 \|g(x)\|_{L^2(\Omega(m))}$$

Consequently, for $n$ sufficiently large,

$$\mathbb{P} \left( \sum_{i=\lceil \sqrt{n} \rceil+1}^{\infty} P^i g(X_k) \right) > \frac{\sqrt{Cn}}{2} \leq \sqrt{D \rho^{\lceil \sqrt{n} \rceil+1} \|g\|_{L^2(\Omega(m))} \leq A \exp(-B \sqrt{n})}$$

for some positive constants $A$ and $B$. Turning to the first term, we have

$$\mathbb{P} \left( \sum_{i=0}^{\lceil \sqrt{n} \rceil} P^i g(X_k) \right) > \frac{\sqrt{Cn}}{2} \leq \sum_{i=0}^{\lceil \sqrt{n} \rceil} \mathbb{P} \left( |P^i g(X_k)|^2 > \frac{Cn}{4} \right),$$

and for $\Omega = \{g^2(X_k) > \frac{n}{4}\} \cup \{g^2(X_k) \leq \frac{n}{4}\}$ gives

$$\mathbb{P} \left( |P^i g(X_k)|^2 > \frac{Cn}{4} \right) \leq \mathbb{P}(g^2(X_k) > \frac{n}{4})$$

$$+ \mathbb{P} \left( \{g^2(X_k) \leq \frac{n}{4}\} \cap \{|P^i g(X_k)|^2 > \frac{Cn}{4}\} \right)$$
and \((iv)\) makes the second term disappear while the first term, thanks to \((ii)\), satisfies

\[
\mathbb{P}\left(g^2(X_k) > \frac{Cn}{4}\right) = \mu \mathbb{P}^k\left(g^2 > \frac{Cn}{4}\right) \leq \exp\left(-\varphi\left(\frac{n}{4}\right)\right).
\]

Combining the above yields

\[
\mathbb{P}\left(\max_{1 \leq k \leq n} |R_k| > \sqrt{\varepsilon Cn}\right) \leq n \mathbb{P}\left(|R_k| > \sqrt{\varepsilon Cn}\right) \leq 2n \mathbb{P}\left(|u(X_k)| > \frac{\sqrt{\varepsilon Cn}}{2}\right) \\
\leq 2n \left[A \exp\left(-B \sqrt{n}\right) + \sqrt{n} \exp(-\varphi\left(\frac{n}{4}\right))\right]
\]

and (4.5.2) follows.

**Theorem 4.12** The following classes of processes satisfy the conclusions (4.5.1) and (4.5.2) of theorem (4.11):

(a) finite state irreducible, aperiodic Markov chains

(b) uniformly ergodic Markov chains with bounded functions \(g\)

(c) Markov chains obtained by quantization \(P^n\) of continuous time Markov processes \(P^t\), that are symmetric on \(L^2(m)\) (i.e. for which \(m\) is a reversible measure).

**Proof:**

(a) Since there is a unique invariant measure \(m = (m_1, ..., m_N)\), the assumption \((i)\) for \(g = (g_1, ..., g_N)\) can be written as \(\sum_{i=1}^{N} g_i m_i = 0\). Then by the Fredholm Alternative, there is a solution of the Poisson equation \((I - P)u \equiv Au = g\) if and only if \(g\) is orthogonal to \(w\) where \(A^*w = A^T w = 0\). The later is true because \(mA = 0\) or \(A^T m^T = 0\). We do not need to verify \((ii)\) since that was only used to prove the existence of solution to the Poisson equation. Also, since \(g\) on \(\{1, ..., N\}\) is bounded, \((iii)\) and \((iv)\) clearly hold.

(b) If the Markov chain is uniformly ergodic then the operator \(Q := I - P\) is normally solvable and condition \((i)\) enables the uniqueness of the solution of
the Poisson equation. That is \( u = R_0 g \) where \( R_0 \) is the potential operator defined by \( R_0 := \sum_{n=0}^{\infty} [P^n - \Pi] \) and the projection operator in \( B(E) \) is defined by \( \Pi g(x) := \int_E m(dx)g(y)\Pi(x) \). Since \( R_0 \) is a bounded operator conclusions (4.5.1) and (4.5.2) follows directly. Moreover, the conditions (iii) and (iv) holds because \( g \) is bounded.

(c) For the sake of the discussion, we consider a class of continuous time Markov processes in \( (\mathbb{R}^d, |\cdot|) \) such that \( P^t g = e^{tA} g \) for smooth functions \( g \), where the infinitesimal generator \( A \) has a discrete spectrum \( \{-\lambda_n, n \in \mathbb{N}\} \) with \( \lambda_0 = 0 < \lambda_1 < \lambda_2 \cdots \) and its corresponding orthonormal in \( L^2(m) \) set of eigenfunctions \( \{e_n, n \in \mathbb{N}\} \) with \( e_0 = 1 \). Then for any \( g \) in the orthogonal complement \( \perp_1 \), the condition (i) is satisfied and \( P^k e_n = e^{-\lambda_n k} e_n \), \( n \geq 1 \). Consequently, for \( g \in \perp_1 \), \( g = \sum_{n=1}^{\infty} \alpha_n e_n \), \( P^k g(x) = \sum_{n=1}^{\infty} \alpha_n e^{-\lambda_n k} e_n \), whence \( \|P^k g(x)\|_{L^2(m)} \leq \rho^k \|g\|_{L^2(m)} \) with \( 0 < \rho = e^{-\lambda_1} < 1 \) and this gives (ii). Furthermore, the invariant measure \( m \) often has a density with respect to Lebesgue measure and condition \( \frac{d\mu P^k}{dm} \leq D < \infty \) is satisfied. Finally, given \( m \), the tail condition \( \int_{\{g^2 > n\}} g^2 \, dm \leq \exp(-\varphi(n)) \) is readily verifiable for a large class of \( g \in \perp_1 \) in addition to which \( g \) is chosen to satisfy (iv). We remark that for any finite linear combination of eigenfunctions from \( \perp_1 \), the condition (iv) is redundant and the tail condition reduces to \( \int_{\{e_i^2 > a\}} e_i^2 \, dm \leq e^{-\varphi(a)} \) for large \( a \). Namely, given \( g = \sum_{i=1}^{N} \alpha_i e_i \),

\[
|P^i g(x)|^2 \leq \|g\|_{L^2(m)}^2 (e_1^2(x) + \cdots + e_N^2(x))
\]

and therefore

\[
\mathbb{P}(|P^i g(X_k)|^2 > n) \leq \mathbb{P} \left( \sum_{i=1}^{N} e_i^2(X_k) > n \right) \leq
\]
\[ \sum_{i=1}^{N} \mathbb{P} \left( e_i^2(X_k) > \frac{n}{N} \right) \leq ND \max_{1 \leq i \leq N} \int_{\{e_i^2 > n/N\}} e^2 dm \]

as claimed.

\[ \square \]

Hypercontractivity Example

Consider a 1-dimensional Ornstein-Uhlenbeck process \( X_t \) satisfying the equation
\[ dX_t = -\frac{1}{2} X_t dt + dB_t, \quad X_0 = x, \]
where \( B_t \) is the standard Brownian motion, with infinitesimal generator
\[ A = \frac{1}{2} \frac{d^2}{dx^2} - \frac{1}{2} x \frac{d}{dx} \]
and the invariant measure \( dm = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx \).

Then \( P^t g(x) := \int g(y)p^t(x, y)dy \), where
\[ p^t(x, y) = \frac{1}{\sqrt{2\pi(1 - e^{-t})}} e^{-\frac{(xe^{\frac{1}{2}} - y)^2}{2(1 - e^{-t})}}, \]
is a \( m \)-symmetric hypercontractive Hermite semigroup with other examples, including Laguerre semigroup, studied in [30] and [31].

Here, for Hermite polynomials \( H_n = (-1)^n e^{-\frac{x^2}{2}} \frac{d^n}{dx^n} e^{-\frac{x^2}{2}} \), we have \( H_0 = 1, \ H_1 = x, \ H_2 = x^2 - 1, \ H_3 = x^3 - 3x, ..., \ e_n = \frac{H_n}{\sqrt{n!}} \) form an orthonormal basis for \( L^2(m) \) and \( AH_n = -\frac{n}{2} H_n \) gives the corresponding eigenvalues \( \lambda_n = -\frac{n}{2}, \ n \in \mathbb{N} \). Letting \( t = 0, 1, 2, ... \) we obtain the embedded Markov chain \( X_n \). Since
\[ d\mu P^k = \frac{1}{\sqrt{2\pi(1 - e^{-k})}} e^{-\frac{(xe^{\frac{1}{2}} - y)^2}{2(1 - e^{-k})}} dy \]
then, for example, taking \( \mu = \delta_x = \delta_0 \) we have
\[ \frac{d\mu P^k}{dm} = \frac{1}{\sqrt{1 - e^{-k}}} e^{-\frac{x^2}{2(1 - e^{-k})}} \leq \sqrt{\frac{e}{e - 1}} \equiv D \]
It is easy to see that for polynomial domination, \( |g(x)| \leq B|x|^l, \ l \in \mathbb{N}, \) and for large \( n \)

\[
\int_{\{g^2 > n\}} g^2 \, dm \leq C n^{2l-1} e^{-\frac{n^2}{2}} \leq e^{-\frac{n^2}{4}}
\]

whence one may choose \( \varphi(x) = \frac{x^2}{4} \) to satisfy the tail estimate.

For example, taking \( g(x) = x \) one gets

\[
|P^n g(x)|^2 = \left( \int |g(y)| p^n(x, y) \, dy \right)^2 \leq D \|g\|_{L^2(m)} = D^2 (1 - e^{-i} + x^2)
\]

\[
\leq 4D \max(1, x^2)
\]

and condition \((iv)\) holds true with \( C = 2D. \)

For general \( g, \) even growing exponentially fast to infinity, it requires some work through estimates and usually \( g \) determines \( \varphi. \)

### 4.6 Large deviation principle for additive functionals

Let \( X_n, \ n \geq 0 \) be an ergodic \( E \)-valued Markov chain with stationary measure \( m, \) as in Theorem 4.11. Let \( S_n = \sum_{k=0}^{n-1} g(X_k) \) and \( \mathbb{E}_m(S_1^2) = \sigma^2 \in (0, \infty). \) According to Theorem 4.11, \( S_n = M_n + R_n, \) where \( M_n \) is a mean zero martingale and \( R_n \) satisfies \((4.5.2). \) Let \( W_n \) and \( W_n^M \) be the empirical measures defined in \((4.1.4) \) and \((4.1.5) \) corresponding to the interpolation processes \( \Psi_n \) and \( \Psi_n^M \) defined by \((4.1.3) \) and \((4.1.6) \) respectively. By Theorem 4.2, \( W_n \) converges weakly to the Wiener measure \( W \) on \( C[0, \infty). \) To conclude our analysis we invoke a martingale LDP \([32], \) applied here to \( W_n^M, \) and show that \( W_n^M \) and \( W_n \) are LDP equivalent.

**Lemma 4.13** If

\[
\lim_{n \to \infty} \frac{1}{\log n} \log \mathbb{P}\{ |\Psi_n - \Psi_n^M|_\phi > \epsilon \} = -\infty \tag{4.6.1}
\]
then \((W_n) \equiv \left( \frac{1}{L(n)} \sum_{k=1}^{n} \frac{1}{k} \delta_{\Psi_k} \right)\) and \((W_n^M) \equiv \left( \frac{1}{L(n)} \sum_{k=1}^{n} \frac{1}{k} \delta_{\Psi_k^M} \right)\) are equivalent with respect to LDP.

**Proof:** By the above Lemma, we need to verify

\[
\lim_{n \to \infty} \frac{1}{\log n} \log \mathbb{P}\{d_\phi(W_n, W_n^M) > \varepsilon\} = -\infty \quad (4.6.2)
\]

where \(d_\phi\) is defined in (4.4.2). For \(f \in \mathcal{C}(\mathcal{C}_\phi, \mathbb{R})\), \(\|f\|_L \leq \frac{1}{2}\) we have

\[
\left| \int f \, dW_n - \int f \, dW_n^M \right| \leq \frac{1}{L(n)} \sum_{k=1}^{n} \frac{1}{k} |f(\Psi_k) - f(\Psi_k^M)| = \\
\frac{1}{L(n)} \sum_{k=1}^{[n^{\varepsilon/2}]} \frac{1}{k} |f(\Psi_k) - f(\Psi_k^M)| + \frac{1}{L(n)} \sum_{k=[n^{\varepsilon/2}]+1}^{n} \frac{1}{k} |f(\Psi_k) - f(\Psi_k^M)|
\]

\[
|\Psi_k - \Psi_k^M|_\phi \leq \frac{L([n^{\varepsilon/2}]/L(n))}{2L(n)} \sup_{1+[n^{\varepsilon/2}] \leq k \leq n} |\Psi_k - \Psi_k^M|_\phi
\]

whence

\[
\mathbb{P}\{d_\phi(W_n, W_n^M) > \varepsilon\} \leq \mathbb{P}\{ \sup_{1+[n^{\varepsilon/2}] \leq k \leq n} |\Psi_k - \Psi_k^M|_\phi > \varepsilon\}.
\]

Condition (4.6.1) implies that for every \(\varepsilon > 0\) and \(N > 0\), there exists \(k_0\) such that for any \(k \geq k_0\), \(|\Psi_k - \Psi_k^M|_\phi < k^{-2N} k^{-2}\). Therefore,

\[
\mathbb{P}\{d_\phi(W_n, W_n^M) > \varepsilon\} \leq \sum_{k=[n^{\varepsilon/2}]+1}^{n} \mathbb{P}\{|\Psi_k - \Psi_k^M|_\phi > \varepsilon\} \leq \\
\sum_{k=[n^{\varepsilon/2}]+1}^{n} k^{-2N} k^{-2} \leq n^{-N} \sum_{k=1}^{\infty} k^{-2} = cn^{-N}.
\]

where \(c\) is a positive constant and (4.6.2) holds.
Assume in addition to the conditions of Theorem 4.11 that
\[ \sum_{k=n}^{\infty} \exp(-\varphi(k)) < \frac{1}{n^{\gamma}}, \text{ for any positive } \gamma, \text{ for large } n \text{ (for example, } \varphi(x) \sim x^{\alpha}, \alpha > 0). \]

**Theorem 4.14** The sequence \((W_n)\) satisfies the large deviation principle with constants \((\log n)\) and rate function \(I|_{\mathcal{M}_1(C_\phi)}\) defined in (4.3.3), that is, for any Borel set \(A \subseteq \mathcal{M}_1(C_\phi)\),

\[
-\inf_{A^c} I \leq \lim \inf_{n \to \infty} \frac{1}{\log n} \log \mathbb{P}\{W_n \in A\} \\
\leq \lim \sup_{n \to \infty} \frac{1}{\log n} \log \mathbb{P}\{W_n \in A\} \leq -\inf_{A} I
\]

where \(\mathcal{M}_1(C_\phi) := \{Q \in \mathcal{M}_1(\mathcal{C}[0, \infty)) : Q(C_\phi) = 1\}\), and \(\phi\) is defined in (4.3.1).

**Proof:**

By Lemma 4.13 it suffices to check that (4.6.1) holds. We have
CHAPTER 5
FURTHER RESEARCH DIRECTIONS

Large deviation principle for additive functionals of Markov chains was determined in Chapter 4. A similar result can be found for additive functionals of Markov processes.

Let \( X_t = \int_0^t g(Y_s) \, ds \) where \( \{Y_t\} \) is a Markov process with invariant measure \( m \) and \( g \in L^2(m) \) such that \( \int g \, dm = 0 \). To \( X_n(t) = \frac{1}{\sigma \sqrt{n}} \int_0^n g(Y_s) \, ds \) we associate the empirical processes \( W_n : \mathcal{C}[0, \infty) \to \mathcal{M}_1(\mathcal{C}[0, \infty)) \),

\[
W_n(\cdot) := \frac{1}{L(n)} \sum_{k=1}^n \frac{1}{k} \delta_{\{X_k \in \cdot\}}
\]

with \( L(n) = \sum_{k=1}^n \frac{1}{k} \).

**Theorem 5.1 (Bhattacharya) (Functional central limit theorem)** Let \( \{Y_t\}_{t \geq 0} \) be an ergodic stationary Markov process on a state space \( S \). If \( A \) is its infinitesimal generator on \( L^2(S, dm) \) where \( m \) is the invariant probability measure, then for all functions \( g \) in the range of \( A \), \( X_n(t) = \frac{1}{\sigma \sqrt{n}} \int_0^n g(Y_s) \, ds, t \geq 0 \), converges to \( W \), the Wiener measure with zero drift and variance parameter \( \sigma^2 = -2\langle g, f \rangle = 2\langle Af, f \rangle \) where \( f \) is some element in the domain of \( A \) such that the Poisson equation \( g = -Af \) holds.

This theorem can be easily extended to \([0, \infty)\).

In order to establish the functional almost everywhere convergence (FAECLT) of \( W_n \) to a Wiener measure on \( \mathcal{C}[0, \infty) \) we need to find a martingale decomposition for the additive functional \( X_n(t) \).

Using Poisson equation and the martingale problem associated to the Markov process \( Y_t \), one gets
\[ X_n(t) = \frac{1}{\sigma \sqrt{n}} \int_0^{nt} g(Y_s) \, ds = -\frac{1}{\sigma \sqrt{n}} \int_0^{nt} A f(Y_s) \, ds \]
\[ = f(Y_t) - f(Y_0) - \frac{1}{\sigma \sqrt{n}} \int_0^{nt} A f(Y_s) \, ds + (f(Y_0) - f(Y_t)) \]
\[ = M_n(t) + R_n(t) \]

where \( M_n(t) \) is a martingale. Next it should be shown that the remaining part \( R_n(t) \) goes to 0 in probability and the FAECLT follows. For the large deviation result we have to prove that \( R_n(t) \) goes to 0 to a certain rate, and we need to find criteria for which this is fulfilled.

Another future direction is to find a large deviation result for the stochastic additive functional

\[ \xi^\varepsilon(t) = \xi^\varepsilon(0) + \int_0^t \eta^\varepsilon(ds; x^\varepsilon(s;
\frac{s}{\varepsilon^2})). \]

where the supporting Markov process \( x(t), t \geq 0 \) is uniformly ergodic on \( E \) with the stationary distribution \( \pi(dx) \), and \( \eta^\varepsilon(t;x), t \geq 0, x \in E \) is a process with locally independent increments defined by the generators

\[ \Pi^\varepsilon(x)\varphi(u) = a^\varepsilon(u;x)\varphi'(u) + \varepsilon^{-1} \int_{\mathbb{R}^d} [\varphi(u + \varepsilon v) - \varphi(u) - \varepsilon v \varphi'(u)] \Gamma^\varepsilon(u, dv; x). \]

To get a martingale decomposition for this stochastic additive functional one can consider the Poisson equation associated with it. It could be shown that the existence of the solution for the Poisson equation of the process with locally independent increments is equivalent to the existence of the solution of the Poisson equation for the embedded Markov chain associated with the process with locally independent increments. We note that because of the random jumps of the process, the Poisson equation for continuous time processes cannot be applied directly. Using a similar
approach as in Chapter 4 one can get the large deviation result for the stochastic additive functional $\xi^\epsilon(t)$.

Another future research goal is that of extending this dissertation to limit theorems and large deviations for stochastic additive functionals switched by semi-Markov processes. Exit time applications and stability of perturbed dynamical systems may also be considered.
REFERENCES


BIOGRAPHICAL INFORMATION

Adina Oprisan was born in Campulung, Romania in 1974. She graduated from Bucharest University, Romania in 1997, after which she received a M.S. in Mathematics from the same university in 1998. From 1998 to 2003 she was with the University “Politehnica” of Bucharest as an instructor in Department of Mathematics II. She received a M.S. degree in Statistics from Michigan State University in 2006 and her Ph.D. in Mathematics from the University of Texas at Arlington in 2009. Her current research interests lie in the area of stochastic processes, with focus on randomly perturbed dynamical systems, almost everywhere central limit theorems, large deviation principle. She is a member of AMS and SIAM.